理學博士學位論文

Theoretical Study on Conformation-Dependent Properties of the Three Aromatic Amino Acids

세 가지 방향족 아미노산의 구조-의존적 특성에 관한 이론적 연구

2014 年 2 月

서울大學校 大學院
化學部 物理化學專攻
白 炅 燁

Ph.D. Dissertation

Theoretical Study on Conformation-Dependent Properties of the Three Aromatic Amino Acids

February 2014

Supervisor: Professor Seong Keun Kim

Co-supervisors: Professor Dr. Sheng Hsien Lin

Professor Dr. Yuichi Fujimura

Department of Chemistry
Seoul National University
Kyung Yup Baek

Abstract

Theoretical Study on Conformation-Dependent Properties of the Three Aromatic Amino Acids

Kyung Yup Baek

Department of Chemistry

The Graduate School

Seoul National University

Study on conformation-dependent properties of the three aromatic amino acids (L-phenylalanine, L-tyrosine and L-tryptophan) in neutral and radical cations has been performed by using density functional theory (DFT) with a new density functional M05-2X, which is applicable to molecular systems with noncovalent long-range interactions such as intramolecular hydrogen bonding.

First of all, M05-2X is applied to several simple chromophores in order to verify reliability of the method as well. The results obtained from M05-2X are compared with those estimated by using the conventional DFT (B3LYP) and

can explain the correspondence between the observed and predicted chromophores without ambiguity. On the other hand, a considerable discrepancy between the predicted results from DFT (B3LYP) and experimental ones is found. This suggests that noncovalent long-range effects should be included for accuracy of calculation.

Optimized geometrical structures for both the neutral and cationic conformers of the amino acids are evaluated with those DFT methods and the procedure of conformerization for L-phenylalanine conformers in neutral is indicated from the results of intrinsic reaction coordinate (IRC) profile.

As described above, M05-2X method successfully reproduces experimental adiabatic and vertical ionization energies for all the three aromatic amino acids, whereas B3LYP functional yields significantly lower ionization energies from the observed ones by 0.2 ~ 0.3 eV.

Charge distributions in the cationic conformers are estimated by using natural bond orbital (NBO) analysis. All the conformers of L-tyrosine and L-tryptophan become charge localization when they are ionized regardless of the type of intramolecular hydrogen bonding, unlike the case of L-phenylalanine that was clarified earlier by other studies.

Their charge distribution analysis helps dynamics of charges in proteins to be understood because the three aromatic amino acids play a role as chromophores in biomolecules.

A simple model is employed in order to figure out different degrees of charge localization among all the three aromatic amino acids. The basic concept of the model starts from postulation that the aromatic amino acid consists of two sub-moieties of distinct cationic core: the backbone and aromatic residue. The difference in adiabatic ionization energy between these two sub-moieties is found to govern the degree of charge localization.

Here it is illuminated theoretically that there exists considerable correlation between conformation-dependent property and ionization energy for the three aromatic amino acids.

Keywords: M05-2X, noncovalent long-range effect, L-phenylalanine, L-tyrosine and L-tryptophan, charge localization, ionization energy

Student Number: 2002-30939

Contents

Chapter 1.

Introduction

In 1900, Max Planck offered a radical proposal that blackbody radiation emitted by microscopic particles was limited to certain discrete values, i.e., it was 'quantized'.[1] Such quantization was essential to reconciling large differences between predictions from classical models and experimental results.

As the twentieth century progressed, it became increasingly clear that quantization was not only a characteristic of light, but also of the fundamental particles from which matter is constructed. Bound electrons in atoms, in particular, are clearly limited to discrete energies (levels) as indicated by their ultraviolet and visible line spectra.

This phenomenon has no classical correspondence – in a classical system, obeying Newtonian mechanics, energy can vary continuously. In order to describe microscopic systems, another different mechanics was required.

Matter at the atomic and molecular or microscopic level reveals the existence of a variety of particles which are identifiable by their distinct properties, such as mass, charge, spin, and magnetic moment. All of these seem to be of a quantum nature in the sense that they take on only certain discrete values. This discreteness of physical properties persists when particles combine to form nuclei, atoms, and molecules.[2]

The development of quantum mechanical techniques are more generally applicable and can be implemented on a computer. Now, quantum mechanics can be used to perform calculations molecular systems of real and practical interest.[3]

Computational chemistry which attained general currency as the same meaning of theoretical chemistry is based on molecular quantum mechanics with Schrödinger equation. It has been great tool which is employed to understand molecular properties to the motion and interactions of electrons and nuclei.[4]

Actually, any chemical species (unstable conformers and other reactive intermediates in neutral) which might be difficult to investigate experimentally may be predicted theoretically.[5]

Over the last century, there have been accelerated advances in the field of computational chemistry and they have enabled development of efficient quantum chemical calculation methods employing electronic structure theories such as DFT (B3LYP)[6-8] and a new DFT functional M05-2X.[9,10]

These theoretical methods are essential for many studies on optimized geometry and electronic properties. Nowadays, combinations of supercomputers with high-performance and Gaussian software package which is one of enhanced electronic structure programs are extremely useful to interpret complicated experimental data.

1.1. The three aromatic amino acids

Because of importance as building blocks of proteins and roles in biomolecules, many studies on amino acids have been done through laser spectroscopic techniques combined with quantum chemical calculation methods.[5,11-19]

The three aromatic amino acids (Figure 1.1) can be divided into two substructure, backbone of α-amino acid like alanine and aromatic residue like toluene, cresol and skatol. L-phenylalanine and L-tyrosine have same structures except –OH group attached to aromatic ring.

It has been shown that the aromatic amino acids have many conformers depending on internal charge distributions and external conditions.[5,11-14] It was reported that L-phenylalanine is biologically converted into L-tyrosine in the liver by phenylalanine hydroxylase and L-tyrosine is a precursor to neurotransmitter dopamine which is known to be related to Parkinson's disease.[20,21]

On the other hands, L-tryptophan is known to be a precursor for serotonin which is a neurotransmitter modulating central and peripheral functions through action on neurons, smooth muscle, and other cell types.[22,23]

It was reported that tryptophan depletion consequently brings to significant reduction serotonin (5-HT) synthesis and release in brain and lack of serotonin recurs symptoms in patients with depression and panic disorder, who have responded to treatment with antidepressant due to reduction.[24]

Since chromophores of these three aromatic amino acids have high optical absorption cross sections at 250 – 285 nm in the ultraviolet (UV) region,[25] their spectroscopic studies have been used as optical probes for structures and dynamics of proteins.[26]

Aromatic amino acids have also been utilized as a trigger to charge transfer dynamics of polypeptides.[14,27,28]

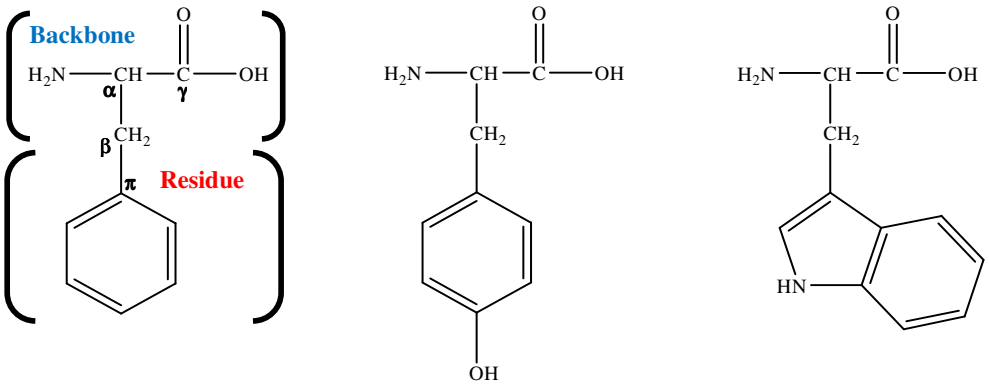


Figure 1.1. Molecular structure of L-phenylalanine, L-tyrosine and L-tryptophan.

1.2. Intramolecular interactions

A hydrogen bond is a kind of typical noncovalent bond, that is, it is not direct covalent bond but electromagnetic attractive interaction between polar atoms (molecules). This interaction is called “intermolecular” hydrogen bonding when it is occurred between molecules, whereas it is called “intramolecular” hydrogen bonding when it is occurred within different parts of a single molecule.

According to types of intramolecular hydrogen bonding, all conformers of the aromatic amino acids were known to be classified into two subgroups, I and II.[29] (Subgroup I characterized by $-COOH \rightarrow -NH_2$ vs. subgroup II by $-NH_2 \rightarrow -OCOH$ in Figure 1.2.)

Noncovalent long-range interactions between a partially charged backbone and a residue in the amino acids play a crucial role in determining stable conformers, especially, ionic conformers.[14] B3LYP has been considered one of the most famous functional. However, it involves some critical defects due to absence of noncovalent interactions in the functional.[30]

Recently, new functionals including the noncovalent effect have been developed in order to complement the shortcoming of B3LYP, which show good agreement with experiments than B3LYP.[30-36]

For some chromophores, we carried out quantum chemical calculation of geometry optimization and estimated their adiabatic ionization energies based on M05-2X in order to investigate how much long-range effects impact on the accuracy of the results.

Especially, the M05-2X results of 2-phenylethylamine (2-PEA) and L-phenylalanine with intramoelcular hydrogen bonding show good agreement with the experimental data compared with the B3LYP results as shown in Figure 1.3.

This indicates that it is essential to apply a new DFT functional M05-2X to neutrals and, in particular, cationic conformers of amino acids in order to ascertain correspondence between observed and predicted conformers.

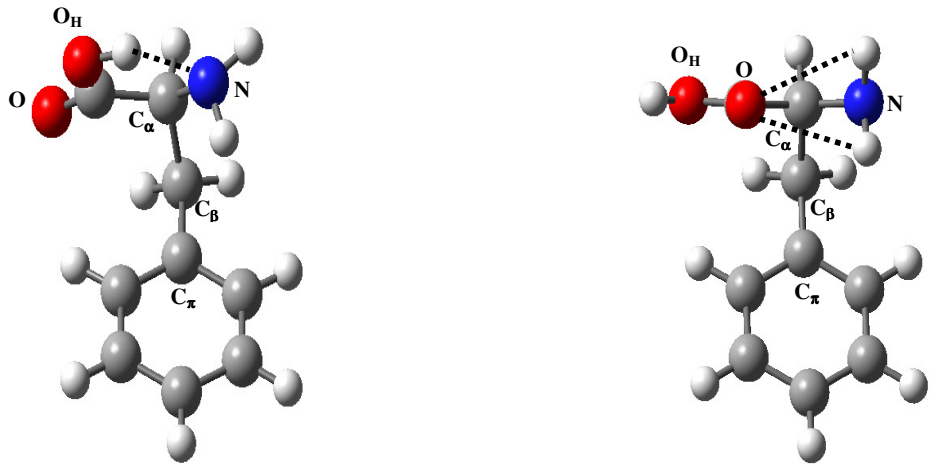


Figure 1.2. Representative structure of subgroup I and II in aromatic amino acid

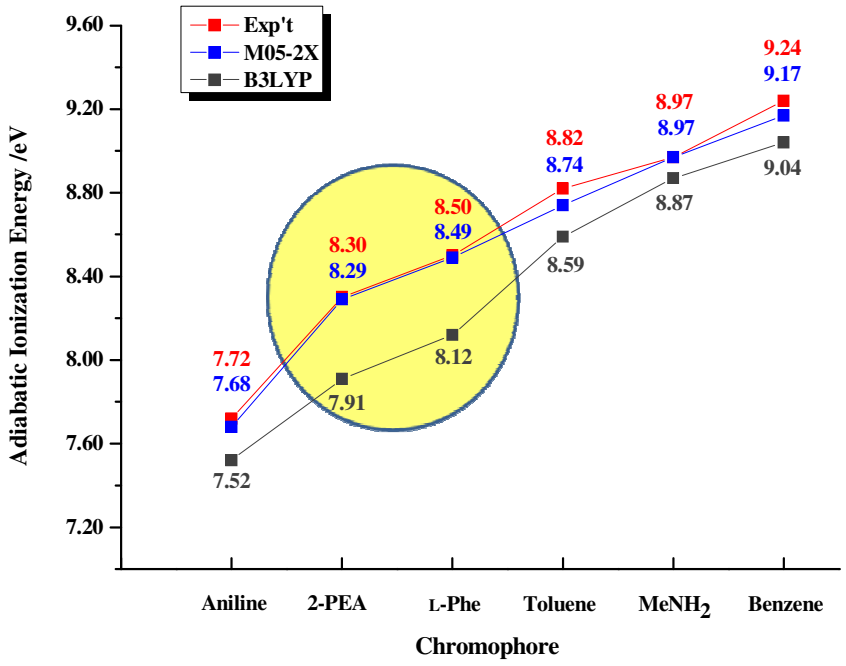


Figure 1.3. Adiabatic ionization energies for several chromophores.

1.3. Charge localization vs. delocalization

Let us consider ionized species of a molecule that consists of two moieties such as residue (A) and backbone (B), respectively as shown in Figure 1.4. There exist two cases depending on whether moieties have same ionization energy or not.

It has been reported that degree of charge localization is proportional to ionization energy gap between two moieties of molecules and 2-PEA, which consists of two moieties (ethylamine and ethylbenzene), were employed as a reference molecule to explain charge delocalization due to be equal to ionization energies each other.[11,37]

This approach is useful to interpret the different tendencies of charge distribution (localization or delocalization) for the three aromatic amino acids (L-phenylalanine, L-tyrosine, and L-tryptophan), and to investigate conformation-dependent properties between the subgroups (subgroup I and II) even though they belong to the same molecule.[5,11]

In the thesis, main topic is M05-2X results on conformation-dependent properties such as geometrical structure, adiabatic (vertical) ionization energies, charge distributions for the neutral (cationic) conformers of these aromatic amino acids. In addition, a simple molecular orbital (MO) model is introduced to explain the different tendency of their ionization energies and charge distributions.

In the next two chapters, a theoretical background relevant to all these work and computational method we have employed is presented.

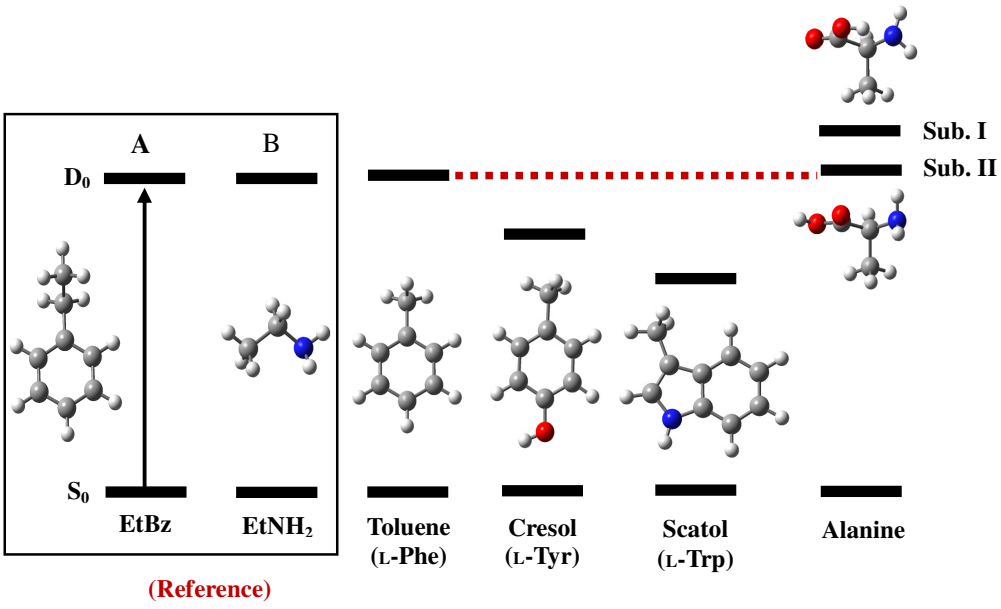


Figure 1.4. Ionized backbones and residues corresponding aromatic amino acids.

1.4. References

[1] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), p. 95.

[2] Eugen Merzbacher, *Quantum Mechanics* (John Wiley & Sons, Ltd., New York, 1998), p. 1.

[3] Andrew R. Leach, *Molecular Modelling* (Anddison Wesley Longman Ltd, Harlow Essex, 1996), p. 25.

[4] Warren J. Hehre; Leo Radom.; Paul v. R. Schleer; John A. Pople, *Ab initio Molecular Orbital Theory* (John wiley & Sons Inc., New York, 1986), p. 1.

[5] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583.

[6] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

[7] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[8] Lee, C.; Yang, W.; Parr, R.G. *Phys. Rev. B* **1988**, *37*, 785.

[9] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[10] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[11] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2011**, *115*, 9658.

[12] Lee, K. T.; Kim, H. M.; Han, K. Y.; Sung, J.; Lee, K. J.; Kim, S. K. *J. Am. Chem. Soc.* **2007**, *129*, 2588.

[13] Lee, K. T.; Sung, J.; Lee, K. J.; Kim, S. K.; Park, Y. D. *Chem. Phys. Lett.* **2003**, *368,* 262.

[14] Weinkauf, R.; Schanen, P.; Yang, D.; Soukara, S.; Schlag, E. W. *J. Phys. Chem.* **1995**, *99*, 11255.

[15] Robertson, E. G.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3,* 1.

[16] Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321,* 49.

[17] Hashimoto, T.; Takasu, Y.; Yamada, Y.; Ebata, T. *Chem. Phys. Lett.* **2006**, *421,* 227.

[18] Lee, Y.; Jung, J.; Kim, B.; Butz, P.; Snoek, L. C.; Kroemer, R. T.; Simons, J. P. *J. Phys. Chem. A* **2004**, *108,* 69.

[19] Huang, Z.; Yu, W.; Lin, Z., *Theochem-J. Mol. Struct.* **2006**, *758,* 195.

[20] Elsworth, J. D.; Roth, H. R. *Exp. Neurol.* **1997**, *144*, 4.

[21] Grace, L. I.; Cohen, R.; Dunn, T. M.; Lubman, D. M.; de Vries, M. S. *J. Mol. Spectrosc.* **2002**, 2*15*, 204.

[22] Shi, Z.; Devasagayaraj, A.; Gu, K.; Jin, H.; Marinelli, B.; Samala, L.; Scott, S.; Stouch, T.; Tunoori, A.; Wang, Y.; Zang, Y.; Zhang, C.; Kimball, S. D.; Main, A. J.; Yu, X.; Buxton, E.; Patel, S.; Nguyen, N.; Swaffield, J.; Powell, D. R.; Wilson, A.; Liu, Q. *J. Med. Chem.* **2008**, *51*, 3684.

[23] Lima, S.; Kumar, S.; Gawandi, V.; Momany, C.; Phillips, R. S. *J. Med. Chem.* **2009**, *52*, 389.

[24] Bell, C.; Abrams, J.; Nutt, D. *Brit. J. Psychiat.* **2001**, *178*, 399.

[25] Steiner, R. F.; Weinryb, I. *Excited States of Proteins and Nucleic Acids* (Macmillan, London, 1971).

[26] Beechem, J. M.; Brand, L. *Annu. Rev. Biochem.* **1985**, *54*, 43.

[27] Weinkauf, R.; Schanen, P.; Metsala, A.; Schlag, E. W.; Burgle, M.; Kessler, H. *J. Phys. Chem.* **1996**, *100*, 18567.

[28] Lehr, L.; Horneff, T.; Weinkauf, R.; Schlag, E. W. *J. Phys. Chem. A* **2005**, *109*, 8074.

[29] Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321,* 49.

[30] Zhao, Y.; Truhlar, D. G. *Accounts Chem. Res.* **2008**, *41,* 157.

[31] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[32] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[33] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123,* 161103.

[34] Suresh, C. H.; Mohan, N.; Vijayalakshmi, K. P.; George, R.; Mathew, J. M. *J. Comput. Chem.* **2008**, *30,* 1392.

[35] Sato, T.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2007,** *126,* 234114.

[36] Hohenstein, E. G.; Chill, S. T.; Sherrill, C. D. *J. Chem. Theory Comput.* **2008,** *4,* 1996.

[37] Weinkauf, R.; Lehrer, F.; Schlag, E. W.; Metsala, A. *Faraday Discuss.* **2000**, *115*, 363.

Chapter 2.

Background

2.1. Schrödinger equation

The fundamental postulation of quantum mechanics is that a so-called wave function, Ψ, exists for any (chemical) system, and that appropriate operators which act upon Ψ generate the observable properties of the system.

Thus, we write

$$H\Psi = E\Psi\,, \tag{2.1}$$

which is the Schrödinger equation.[1] The operator in Eq. (2.1) that returns the system energy, E as an eigenvalue is called the Hamiltonian operator, H. The typical form of the Hamiltonian operator with which we will be concerned takes into account five contributions to the total energy of a system: the kinetic energies of the electrons and nuclei, the attraction of the electrons to the nuclei, and the interelectronic and internuclear repulsions.

In more complicated situations, e.g., in the presence of an external electric field, in the presence of an external magnetic field, in the event of significant spin-orbit coupling in heavy elements, taking account of relativistic effects, etc., other terms are required in the Hamiltonian. Casting the Hamiltonian into mathematical notation, we have

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i \sum_k \frac{e^2 \mathrm{Z}_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 \mathrm{Z}_k \mathrm{Z}_l}{r_{kl}}\,, \tag{2.2}$$

where i and j run over electrons, k and l run over nuclei, $\hbar$ is Planck's

constant divided by 2π, m_e is the mass of the electron, m_k is the mass of nucleus k, ∇^2 is the Laplacian operator, e is charge on the electron, Z is an atomic number, and r is the distance between particles.

Here, Ψ is a function of $3n$ coordinates where n is the total number of particles (nuclei and electrons), e.g., the x, y, and z Cartesian coordinates specific to each particle. If we work in Cartesian coordinates, the Laplacian has the form

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} . \tag{2.3}$$

The Hamiltonian operator in Eq. (2.2) is composed of kinetic and potential energy parts. The kinetic energy for a QM particle, however, is not expressed as $|\mathbf{p}|^2/2m$, but rather as the eigenvalue of the kinetic energy operator

$$T = -\frac{\hbar^2}{2m}\nabla^2. \tag{2.4}$$

According to Born interpretation, the square of an electronic wave function thus gives the electron density at any given point. If we integrate the probability of finding the particle over all space, then the result must be one as the particle must be somewhere:

$$\iiint |\Psi|^2 \, dxdydz = 1 . \tag{2.5}$$

Wave functions which satisfy this condition are said to be normalized. It is usual to require the solutions to the Schrödinger equation to be orthogonal:

$$\iiint \Psi_i^* \Psi_j \, dxdydz = 0 \;\; (i \neq j) \, . \tag{2.6}$$

Here, * denotes complex conjugation of Ψ_j.

A convenient way to express both the orthogonality of the different wave functions and the normalization conditions uses the Kröneckers delta:

$$\int \Psi_i \Psi_j \, d\tau = \delta_{ij} \, . \tag{2.7}$$

Here, $d\tau$ is $3n$-dimensional volume element.

Now, consider the result of taking Eq. (2.1) for a specific Ψ_i, multiplying $\Psi_j$on the left and integrating the resultant equation over τ.

This process gives

$$\int \Psi_j^* H \Psi_i \, d\tau = \int \Psi_j^* E_i \Psi_i \, d\tau \, . \tag{2.8}$$

Since the energy is a scalar value, we can write

$$\int \Psi_j H \Psi_i \, d\tau = E \delta_{ij} \, . \tag{2.9}$$

2.2. The Born-Oppenheimer approximation

Under typical physical conditions, the nuclei of molecular systems are moving far slower than the electrons. Recall the protons and neutrons are about 1800 times more massive than electrons and note the appearance of mass in the denominator of the kinetic energy terms of the Hamiltonian in Eq. (2.2).[1]

As such, it is convenient to decouple these two motions, and compute electronic energies for fixed nuclear positions. That is, the nuclear kinetic energy term is taken to be independent of the electrons, correlation in the attractive electron-nuclear potential energy term is eliminated, and the repulsive nuclear-nuclear potential energy term becomes a simply evaluated constant for a given geometry. Thus, the electronic Schrödinger equation is taken to be

$$(H_{\mathrm{el}} + \mathrm{V}_N)\Psi_{\mathrm{el}}(\mathrm{q}_i; \mathrm{q}_k) = E_{\mathrm{el}}\Psi_{\mathrm{el}}(\mathrm{q}_i; \mathrm{q}_k)\,. \qquad (2.10)$$

H_{el} includes only the first, third, and fourth terms on the right had side of Eq. (2.2), V_N is the nuclear-nuclear repulsion energy, and the electronic coordinates q_i are independent variables but the nuclear coordinates q_k are paprameters (and thus appear following a semicolon rather than comma in the variable list for Ψ).

The eigenvalue of the electronic Schrödinger equation is called the 'electronic energy'. Note that the term V_N is a constant for a given set of fixed nuclear coordinates.

Wave functions are invariant to the appearance of constant terms in the Hamiltonian, so in practice one almost always solves Eq. (2.10) without the inclusion of V_N, in which case the eigenvalue is sometimes called the 'pure electronic energy', and one then adds to this eigenvalue to obtain E_{el}.

2.3. Ionization energy

In 1933, the Dutch theoretical chemist T. Koopmans proposed that ionization energy is equal to the minus of the energy of the orbital (usually HOMO) from which the electron is ejected.[2-4]

The electronic energy of such a system can be calculated in a manner analogous to that for the hydrogen molecule. First, there is the energy of each electron moving in the field of the bare nuclei. For an electron in a molecular orbital χ_i, this contributes an energy H_{ii}^{core}.

If there are two electrons in the orbital then the energy is $2H_{ii}^{core}$ and for $N/2$ orbitals the total contribution to the energy will be[5]

$$\sum_{i=1}^{N/2} 2H_{ii}^{core} . \tag{2.11}$$

The interaction between each pair of orbitals ψ_i and ψ_j involves a total of four electrons if we consider the electron-electron terms. There are four ways in which two electrons in one orbital can interact in a Coulomb sense with two electrons in a second orbital, thus giving $4J_{ij}$. However, there are just two ways to obtain paired electrons from this arrangement, giving a total exchange contribution of $-2K_{ij}$.

Finally, the Coulomb interaction between each pair of electrons in the same orbital must be included; there is no exchange interaction because the electrons

have paired spins. The total electronic energy is thus given as:

$$E = 2\sum_{i=1}^{N/2} H_{ii}^{core} + \sum_{i=1}^{N/2}\sum_{j=i+1}^{N/2} (4J_{ij} - 2K_{ij}) + \sum_{i=1}^{N/2} 2J_{ij} \,. \tag{2.12}$$

A more concise form of this equation can be obtained if we recognize that $J_{ii} = K_{ii}$:

$$E = 2\sum_{i=1}^{N/2} H_{ii}^{core} + \sum_{i=1}^{N/2}\sum_{j=1}^{N/2} (2J_{ij} - K_{ij}) \,. \tag{2.13}$$

A Hartree-Fock calculation provides a set of orbital energies, ε_i. The energy of an electron in a spin orbital is calculated by adding the core interaction H_{ii}^{core} to the Coulomb and exchange interactions with the other electrons in the system:

$$\varepsilon_i = H_{ii}^{core} + \sum_{j=1}^{N/2} (2J_{ij} - K_{ij}) \,. \tag{2.14}$$

where, J_{ij} is often used to represent this Coulmb interation beween electrons in spin orbitals i and j. The energy due to exchange 'interaction' is often represented as K_{ij}. In a closed-shell system containing N electrons in $N/2$ orbials, there are two spin orbitals associated with each spatial orbital ψ_i : $\psi_i\alpha$ and $\psi_i\beta$.

One should be careful when applying Koopman's theorem and comparing the results with experimentally determined ionization energies. It is that the orbitals in the ionized state are assumed to be the same as in the unionized state; they are 'frozen'. This neglects the fact that the orbitals in the ionized state will be

different from those in the unionized state. The energy of the ionized state will thus tend to be higher than it should be, giving too large an ionization energies.[2,4]

2.4. Charge distribution

Charge density[3] at a point in an atom, molecule, or sold is the charge per unit volume at each location. For a one-electron system, it is equal to the absolute square of the wave function at the point of interest multiplied by the electronic charge, $-e$.

For a many-electron system, the total charge density is the sum of all the one-electron contributions. The product of the charge density at a location and the volume element $d\tau$ is equal to the charge inside the volume element.

In Cartesian coordinates, the expectation values of electric multipole moment operators are computed as[6]

$$\langle \mathbf{x}^k \mathbf{y}^l \mathbf{z}^m \rangle = \sum_i^{\text{atoms}} \mathrm{Z}_i x_i^k y_i^l z_i^m - \int \Psi^*(\boldsymbol{\tau}) \left(\sum_j^{\text{electrons}} x_j^k y_j^l z_j^m \right) \Psi(\boldsymbol{\tau}) d\boldsymbol{\tau} \quad , \qquad (2.15)$$

where the sum of $k, l,$ and m determines the type of moment (0 = monopole, 1 = dipole, 2 = quadrupole, etc.), Z_i is the nuclear charge on atom i, and the intergration variable $\boldsymbol{\tau}$ contains the x, y, and z coordinates of all of the electrons j.

The simplest moment to evaluate is the monopole moment, which has only the component $k = l = m = 0$, so that the operator becomes $\mathbf{1}$ and independent of coordinate system we have

$$\langle \mathbf{1} \rangle = \sum_{i}^{\text{atoms}} \mathrm{Z}_i - \sum_{j}^{\text{electrons}} \int \Psi_j(\boldsymbol{\tau}_j)\Psi_j(\boldsymbol{\tau}_j) d\boldsymbol{\tau}_j = \sum_{i}^{\textbf{atoms}} \mathrm{Z}_i - N \ , \qquad (2.16)$$

where N is the total number of electrons.

The monopole moment is thus the different between the sum of the nuclear charges and the number of electrons, i.e., it is the molecular charge. For the dipole moment, there are three possible components: x, y, or z depending on which of k, l, and m is one (with the others set equal to zero).

These are written as μ_x, μ_y, and μ_z. Experimentally, however, one rarely measures the separate components of the dipole moment, but rather the total magnitude, μ, which can be determined as

$$\langle \mu \rangle = \sqrt{\langle \mu_x \rangle^2 + \langle \mu_y \rangle^2 + \langle \mu_z \rangle^2} \ . \qquad (2.17)$$

The dipole moment measures the degree to which positive and negative charge are differential distributed relative to one another, i.e., overall molecular polarity.

2.5. References

[1] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), pp 96-98.

[2] Andrew R. Leach, *Molecular Modelling* (Anddison Wesley Longman Ltd, Harlow Essex, 1996), p. 110.

[3] P.W. Atkins, *Quanta* (Oxford University Press Inc., New York, 1991)

[4] Warren J. Hehre; Leo Radom.; Paul v. R. Schleer; John A. Pople, *Ab initio Molecular Orbital Theory* (John wiley & Sons Inc., New York, 1986), p. 24.

[5] Andrew R. Leach, *Molecular Modelling* (Anddison Wesley Longman Ltd, Harlow Essex, 1996), pp. 53-64.

[6] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), pp 275-276.

Chapter 3.

Computational Methods

3.1. Introduction

In this chapter, theoretical background, methodology, and computational method will be presented. In the first section, an introduction to theoretical background and methodology employed for our study will be indicated briefly.

In the second section, how the calculations were performed will be described in detail.

3.1.1. Density functional theory

The density functional theory (DFT) is widely applied to calculations of the electronic structure of atoms and molecules. Hohenberg and Kohn presented in 1964 that all the ground state properties of a system are functions of the charge density and their theorem is motivation of density functional theory.[1]

It means that an incorrect density gives energy above the true energy. Thus, it is most important to choose the successful candidate functional. The reasonable density will determine suitable wave function and Hamiltonian. That being the case, we can evaluate the energy expectation value

$$\langle \Psi_{\text{candidate}} | H_{\text{candidate}} | \Psi_{\text{candidate}} \rangle = E_{\text{candidate}} \geqq 0 \,. \qquad (3.1)$$

This expression indicates that the expectation value by the variational principle, must be greater than or equal to the true ground state energy.[2] So, in principle, we can select more optimized densities to be closer and correct.

According to the theorem, the total electronic energy should be expressed as a function of the electron density ρ:

$$E(\rho) = E_{KE}(\rho) + E_C(\rho) + E_H(\rho) + E_{XC}(\rho) \ , \qquad (3.2)$$

where $E_{KE}(\rho)$ is the kinetic energy, $E_C(\rho)$ is the electron-nuclear interaction term, $E_H(\rho)$ is the electron-electron Coulombic energy, and $E_{XC}(\rho)$ contains the exchange and correlation contributions.[1]

All of the electron-electron interactions are thus contained within the $E_H(\rho)$ and $E_{XC}(\rho)$ terms. To perform a density functional calculation, it is necessary to rewrite equation (3.2) in terms of the density and then the optimized energy with respect the density, subject to any constraints on the system. The difficulty comes from the electron-electron interaction term in the correct Hamiltonian.[2]

For the exchange-correlation term $E_{XC}[\rho(\mathbf{r})]$, some approximation must be made. The most common way to obtain this contribution uses the so-called local density approximation (LDA), which is based upon a model called the uniform electron gas. In a uniform electron gas, the electron density is constant throughout all space.

The exchange-correlation energy can be determined for this model. In the commonly used Kohn-Sham implementation, the density is represented as if it were derived from a single Slater determinant with orthonormal orbitals.[3]

The various components that contribute to the energy (3.2) must first be expressed in terms of the density. The charge density at a point $\mathbf{r}$ can be written as the sum over occupied molecular orbitals of ψ^2:

$$\rho(\mathbf{r}) = \sum_{\mathrm{i}=1}^{\mathrm{occ}} |\psi_{\mathrm{i}}(\mathbf{r})|^2 \,. \tag{3.3}$$

In ref 3, the use of these 'Kohn-Sham orbitals' with equation (3.3) enables the energy to be optimized by solving a set of one-electron Schrödinger equations as follow

$$\left[-\frac{\nabla^2}{2} + r_{\mathrm{nuclear}}(\mathbf{r}) + \int d\mathbf{r}' \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} + r_{\mathrm{XC}}(\mathbf{r})\right]\psi_{\mathrm{i}}(\mathbf{r}) = \varepsilon_{\mathrm{i}}\psi_{\mathrm{i}}(\mathbf{r}) \tag{3.4}$$

and exchange-correlation functional is the derivative of the exchange-correlation energy with respect to the density:

$$r_{\mathrm{XC}}(\mathbf{r}) = \frac{\delta \mathrm{E}_{\mathrm{XC}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \,. \tag{3.5}$$

The energy is given by

$$E(\rho) = \sum_{\mathrm{i}=1}^{N} \int \psi_{\mathrm{i}} \left(-\frac{\nabla^2}{2}\right) \psi_{\mathrm{i}} d\mathbf{r} + \int r_{\mathrm{nuclear}} \rho(\mathbf{r}) d\mathbf{r}$$

$$+ \frac{1}{2} \iint d\mathbf{r}\, d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \mathrm{E}_{\mathrm{XC}}[\rho(\mathbf{r})] \,. \tag{3.6}$$

DFT calculations requires about the same amount of computation resource as Hatree-Fock theory but they include the effect of electron correlation.[4] This is one of the key advantages of the density functional approach.

3.1.2. Density functionals

(1) B3LYP

Stevens **et al.** modified B3PW91 and developed this method in 1994.[5] The name of functional, B3LYP, implies its use of a three-parameter scheme, as well as the generalized gradient approximation (GGA) exchange and correlation functional B and LYP, respectively.

Let us recall the energy for Hartree-Fock (HF) and density functional theory (DFT). They can be expressed in terms of simplified formula[6] for convenience in equations (3.7) and (3.8), respectively.

$$E_{HF} = \langle h\rho \rangle + V + \frac{1}{2}\langle PJ(\rho)\rangle + E_X[\rho]\,. \qquad (3.7)$$

Here, $\langle h\rho \rangle$ is the one-electron (kinetic + potential) energy, V is the nuclear repulsion energy, $\frac{1}{2}\langle PJ(\rho)\rangle$ is the classical coulomb repulsion of the electrons, and $E_X[\rho]$ is the exchange energy resulting from the quantum nature of electrons.

In the Kohn-Sham formulation of density functional theory, energy E_{KS} is given as,

$$E_{KS} = \langle h\rho \rangle + V + \frac{1}{2}\langle PJ(\rho)\rangle + E_X[\rho] + E_C[\rho]\,. \qquad (3.8)$$

In equation (3.8), $E_C[\rho]$ is the correlation functional. The exact exchange (HF) for a single determinant is replaced by a more general expression, that is, the

exchange-correlation functional, which can include terms accounting for both the exchange and the electron correlation energies, the latter not being present in Hartree-Fock theory as above, i.e., HF theory also can be regarded as a special case of DFT, in which $\mathrm{E_X}[\rho]$ is given by the exchange integral $-\frac{1}{2}\langle \mathrm{PK}(\rho)\rangle$ and $\mathrm{E_C}[\rho] = 0$ within the Kohn-Sham formulation.

A "hybrid" exchange-correlation functional is usually constructed as a linear combination of the HF exact exchange functional, $\mathrm{E_x^{HF}}$ given as,

$$\mathrm{E_X^{HF}} = \frac{1}{2}\sum_{i,j}\iint \psi_i^*(\mathbf{r_1})\psi_j^*(\mathbf{r_1})\frac{1}{\mathbf{r_{12}}}\psi_i(\mathbf{r_2})\psi_j(\mathbf{r_2})d\mathbf{r_1}d\mathbf{r_2}\,, \qquad (3.9)$$

and the parameters determining the weight of each individual functional are typically specified by fitting the functionals predictions to experimental or accurately calculated thermochemical data.

B3LYP is a kind of hybrid density functional theory models to improve calculation accuracy by using correction to exchange-correlation energy and "hybrid" means a mixing other exchange-correlation energy functional in density functional theory (DFT) such as equation (3.10).[5,7-9]

Note the conventional definition of $\mathrm{E_{XC}}$ is equal to sum of exchange energy and correlation energy.

$$E_{\mathrm{XC}}^{hyb} = (1-a)\mathrm{E}_{\mathrm{XC}}^{\mathrm{DFT}} + a\mathrm{E}_{\mathrm{X}}^{\mathrm{HF}} = \mathrm{E}_{\mathrm{XC}}^{\mathrm{DFT}} + a\left(\mathrm{E}_{\mathrm{X}}^{\mathrm{HF}} - \mathrm{E}_{\mathrm{X}}^{\mathrm{DFT}}\right) - a\mathrm{E}_{\mathrm{C}}^{\mathrm{DFT}}, \qquad (3.10)$$

where a is an empirical constant to be optimized by appropriate fits to firm experimental data.

Exchange-correlation functional for B3LYP is

$$E_{\mathrm{XC}}^{\mathrm{B3LYP}} = E_{\mathrm{XC}}^{\mathrm{LSDA}} + \alpha_0\left(E_{\mathrm{X}}^{\mathrm{exact}} - E_{\mathrm{X}}^{\mathrm{LSDA}}\right) + \alpha_{\mathrm{x}}\Delta\mathrm{E}_{\mathrm{X}}^{\mathrm{B}} + \alpha_{\mathrm{c}}\Delta\mathrm{E}_{\mathrm{c}}^{\mathrm{LYP}}, \qquad (3.11)$$

where α_0, α_{x}, and α_{c} are *semiempirical coefficients* to be determined by an appropriate fit to experimental data, $E_{\mathrm{X}}^{\mathrm{exact}}$ is extract exchange energy, $\Delta\mathrm{E}_{\mathrm{X}}^{\mathrm{B}}$ is Becke's gradient correction (to the LSDA) for exchange, and $\Delta\mathrm{E}_{\mathrm{c}}^{\mathrm{LYP}}$is gradient correction for correlation if LYP.[5,7]

Although density functional theory is widely used in the field of computational chemistry, the most popular density functional, B3LYP, has some serious shortcomings:[10] (1) it results in irrelevant for transition metals, (2) it systematically underestimates reaction barrier heights; (3) it is inaccurate for interactions dominated by medium range correlation energy, such as van der Waals attraction, $\pi\cdots\pi$ stacking, and alkane isomerization energies.

The shortcomings originate from the parameter fitting carried out near the region of equilibrium point.

(2) M05-2X

Density functionals using the local spin density approximation (LSDA) depends only on density, however, ones in the generalized gradient approximation (GGA) depends on density and its reduced gradient.[11] GGA functionals have been shown to give more accurate predictions for thermochemistry than LSDA ones, but they still underestimate barrier heights.

Two additional variables, the spin kinetic energy densities are included in the functional form; such functionals are called meta-GGAs. LSDAs, GGAs, and meta-GGAs are "local" functionals because the electronic energy density at a single spatial point depends only on the behavior of the electronic density and kinetic energy.[10, 11]

The long-range correction scheme is inapplicable to most of generalized gradient approximation (GGA) exchange functions. This is because conventional functionals are not usually derived from the corresponding first-order density matrix, although it is necessary to define the dependence of an exchange functional on electron-electron distance.[12]

M05-2X[10,11,13] developed by D. G. Truhlar's group was designed to be suitable for main-group thermochemical kinetics, and noncovalent interactions (weak interaction such as hydrogen bonding, $\pi \cdots \pi$ stacking, interactions energies of nucleobases) by employing a high-nonlocality functional with double the amount of nonlocal exchange (2X).

It adopts hybrid meta-GGA exchange and correlation functionals including all include terms that depend on the kinetic energy density. M05-2X of M05 series was reported to have best performance for weak interactions excluding metal. Showing the detail of the functional, it is composed of four parts: electron spin density, density gradient, kinetic energy density, and Hartree-Fock (HF) exchange.

The Hybrid exchange-correlation energy can be written as follows:[13]

$$E_{\mathrm{XC}}^{\mathrm{hyb}} = E_{\mathrm{X}}^{\mathrm{HF}} + \left(1 - \frac{X}{100}\right)\left(E_{\mathrm{X}}^{\mathrm{DFT}} - E_{\mathrm{X}}^{\mathrm{HF}}\right) + E_{\mathrm{C}}^{\mathrm{DFT}} \quad , \qquad (3.12)$$

where $E_{\mathrm{X}}^{\mathrm{HF}}$ is the nonlocal HF exchange energy, X is the percentage of HF exchange in the hybrid functional, $E_{\mathrm{X}}^{\mathrm{DFT}}$is the local DFT exchange energy, and $E_{\mathrm{C}}^{\mathrm{DFT}}$ is the local DFT correlation energy. In equation (3.12), it is very important that nonlocal HF exchange, $E_{\mathrm{X}}^{\mathrm{HF}}$, be used in M05-2X. On the other hand, that of B3LYP results in density functional theory.

One can see that the total correlation energy for a DFT calculation is modeled as the sum of the dynamic correlation energy given by $E_{\mathrm{C}}^{\mathrm{DFT}}$ and the nondynamic correlation energy contained in

$$\left(1 - \frac{X}{100}\right)\left(E_{\mathrm{X}}^{\mathrm{DFT}} - E_{\mathrm{X}}^{\mathrm{HF}}\right).$$

The M05-2X functional is classified hybrid meta-generalized gradient approximations (hybrid meta-GGAs) because all components except HF exchange are local properties of the density.

Especially, it was reported that the M05-2X give the best performance for noncovalent interactions.[10,13]

3.2. Computational details

All calculations in this thesis were carried out using Gaussian 09 Rev. A01 software package[06] to employ the M05-2X method.[11] The neutral and cationic geometries of conformers of the three aromatic amino acids in the ground state were optimized at M05-2X/6-311+G* level and the same calculation was also carried out with B3LYP/6-311+G* in order to compare with the M05-2X results.

Relative energies between conformers belong to a same aromatic amino acid also were evaluated through the geometry optimization. The optimized structure in corresponding transition state (TS) was obtained for the most probable reaction path between two neutral conformers belonging to the same subgroup, which was confirmed using IRC profiles[06]. Ionization energies for the three aromatic amino acids were also computed with same method by setting the charge and the spin multiplicity to +1 and 2, respectively.

In order to get more comparable results between theoretical estimates and experimental data for all the three aromatic amino acids, zoro-point energy correction[06] was performed together with M05-2X/6-311+G*. Naturally, charge distribution for cations also was obtained using NBO[06] analysis.

The expectation values of the square of the total electron spin, $<S^2>$ was examined in order to check the spin contamination of cationic wave functions. The values of $<S^2>$ calculated by unrestricted M05-2X (B3LYP) were

approximately 0.750 (0.750) for all conformers of cationic three aromatic amino acids. We used Gaussian 09 Rev. A01 installed in the Prof. Hayashi's cluster machine of Taiwan National University. The GaussView 5 (Gaussian Inc.)[15] was employed in order to extract their initial Cartesian coordinates and visualize the result after normal termination for conformers of the three aromatic amino acids, respectively.

Figure 3.1 shows an example of the input file for the Gaussian 09 job and it consists of command lines and information of geometry of molecule[4]. The symbol, % represents command line in the input file. The first command "Nprocshared=" designates the number of processors to use for the Gaussian 09 job. The second one, "Mem=" allocates memory resource for the calculation. The third one, "chk=" saves the calculation result into the predetermined checkpoint file in the command line. The forth command line includes calculation method (M052X), basis set (6-311+G*), calculation purpose (Opt, geometry optimization for the molecule) and so on.

The next section in the input file specifies the charge and the spin multiplicity (0 1), Cartesian coordinate for the molecule. Note insertion of a blank in the last line.

```
%Nprocshared=4
%Mem=8GB
%chk=M052X_1-X.chk
#p Iop(7/8=11) M052X/6-311+G* Opt(Calcall, Maxcyc=300) Pop=NBO

M052X_1-X

 0  1
 C              0.77112500          0.26916200         -0.68296800
 C              0.97200800         -1.09156500         -0.44121800
 C              2.13060700         -1.53137100          0.18628600
 C              3.10559600         -0.62155200          0.58044300
 C              2.91861100          0.73345500          0.34003500
 C              1.75785500          1.17257000         -0.28809900
 H              0.21938800         -1.80376100         -0.75218400
 H              2.27380800         -2.58856000          0.36408600
 H              4.00761100         -0.96788400          1.06617100
 H              3.67616200          1.44740400          0.63403800
 H              1.62172900          2.22908800         -0.48887500
 C             -0.50534100          0.76267600         -1.31615300
 C             -1.65718500          0.85416100         -0.30031900
 H             -0.81784600          0.09781000         -2.11931700
 H             -0.34381000          1.75408500         -1.74118600
```

Figure 3.1. An example of the input file for the Gaussian 09 job.

3.3. References

[1] Andrew R. Leach, *Molecular Modelling* (Anddison Wesley Longman Ltd, Harlow Essex, 1996), p.p. 528-530.

[2] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), pp 233-242.

[3] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

[4] J. B. Foresman and M. J. Frisch, *Exploring Chemistry with Electronic Structure Methods*, 2nd ed. (Gaussian, Inc., Pittsburgh, 1996)

[5] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), pp 248-251.

[6] Gaussian 09, Revision A.1, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.;

Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2009**.

[7] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[8] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

[9] Lee, C.; Yang, W.; Parr, R.G. *Phys. Rev. B* **1988**, *37*, 785.

[10] Zhao, Y.; Truhlar, D. G. *Accounts Chem. Res.* **2008**, *41,* 157.

[11] Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109,* 5656.

[12] Iikura, H.; Tsuneda, T.; Yanai, T.; and Hirao, K. *J. Chem. Phys.* **2001**,115, 3540.

[13] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[14] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583.

[15] GaussView, Version 5, Dennington, Roy; Keith, Todd; Millam, John. Semichem Inc., Shawnee Mission, KS, **2009**.

Chapter 4.

Structure and Energy for the Three Aromatic Amino Acids in Neutral

4.1. Introduction

Assignment of structures between their conformers is not easy as one think in spite of remarkable development of experimental techniques and computational methods. Recently, the trend of preference for calculation method is going to be divided depending on study of purpose in chemistry and physics, that is, the most popular hybrid density functional theory (DFT) like B3LYP can be the best choice if research taget is thermochemistry but it had better to avoid the method if target of study involves weak interaction. [1]

For L-phenylalanine, two conformers of the observed six conformers finally became reassigned and some of the stable conformers predicted by B3LYP calculation could not be observed experimentally.[2-4]

So far, most of the theoretical analyses of conformer spectra have been conducted by using B3LYP excluded long-range effect even though it was certainly known that the three aromatic amino acids (L-phenylalanine, L-tyrosine, and L-tryptophan) are noncovalent long-range interactions such as intramolecular hydrogen bonding.[6-11]

Therefore, it is worth applying a new DFT functional M05-2X to neutrals and cationic conformers of these aromatic amino acids in order to theoretically confirm their assignments and clarify correspondence between observed and predicted conformers. In particular, unlike L-phenylalanine, the other cationic

conformers corresponding to L-tyrosine and L-tryptophan remaining largely unexplored have a value to be investigated. In the next section, theoretical methods and procedures are briefly described.

Section 4.3 provides explicit structural assignment of neutral and cationic conformers of all the three aromatic amino acids based on the M05-2X results of optimized geometrical structure which is a kind of conformation-dependent properties.[12,13] Assignment of some conformers for L-phenylalanine have been argued so far.

Therefore, the most probable reaction path of the argued conformers is theoretically investigated along with intrinsic reaction coordinate (IRC)[14] profile in section 4.3. The results Their optimized structures and energies obtained by B3LYP are also shown in order to compare with that of M05-2X.

4.2. Computational procedure

The each geometry of neutral conformers of L-phenylalanine in the ground state is optimized at the M05-2X/6-311+G* level including zero-point energy correction (ZPE_{corr}) with scanning for five dihedral angles rotation by setting the charge and spin multiplicity to 0 and 1, respectively: C_γ–O = 0°, 180°; C_π–C_β = 60°, 120°,180°; C_β –C_α = -120°, -60°, 0°, 60°, 120°, 180°; C_α–N = -120°, -60°, 0°, 60°, 120°, 180°; C_α–C_γ = 0°, 30°, 90°, 150°, 210°, 270°, 330° in Figure 4.1(a).

A total of 1296 initial geometries of the multiconformer exist, and they are converged to 39 stable local minima and 38 local minima in B3LYP with 6-311+G* and the results obtained from B3LYP reproduce exactly the report by Lin **et al.**[15]

On the other hand, that of cationic conformers of L-phenylalanine in ground state are also optimized at the M05-2X/6-311+G* level including zero-point energy correction (ZPE_{corr}) by setting the charge and spin multiplicity to +1 and 2 using their optimized neutral structures as a initial geometry, respectively.

For L-phenylalanine, the optimized structure in the corresponding transition state (TS) is obtained for the most probable reaction path between two neutral conformers belonging to the same subgroup, which is confirmed using IRC profiles.[14]

In the same manner, the geometries of neutral conformers of L-tyrosine and L-tryptophan in the ground state are optimized at the M05-2X level with the 6-311+G* basis set, based on the similar conformations of the backbone (alanine) in L-phenylalanine.[16]

The eighteen most stable neutral conformers are represented for both L-tyrosine and L-tryptophan by considering the orientation of the phenolic –OH group for L-tyrosine and the asymmetry of the indole ring for L-tryptophan.

In similar of L-phenylalanine, for each conformer of L-tyrosine and L-tryptophan radical cation in the ground state, geometry optimization is calculated at M05-2X level with the 6-311+G* basis set by setting the charge and the spin multiplicity to +1 and 2, respectively.

4.3. Results and discussion

4.3.1. Geometrical structures of neutral, and cationic conformers for the three aromatic amino acids

(1) L-phenylalanine

Figure 4.1a(b) show the nine stable neutral L-phenylalanine conformers and six cationic ones in their ground states. Basically, they were calculated by using M05-2X optimization and conformers obtained by using the popular (DFT) B3LYP, which were different from the M05-2X results, are illustrated together.

A comparison of M05-2X and B3LYP results on their molecular constants for L-phenylalanine conformers is shown in Table 4.1. Here, dihedral angles, N–C_α–C_β–C_π and C_α–NH_2, and bond lengths, C_α–C_β, C_α–N, and C_β–C_π, which characterize each conformer, are listed.

The two functionals, M05-2X and B3LYP, gave almost the same results regarding geometrical structures for other groups such as the phenyl ring and the carboxyl groups in both the neutrals and the cations.

The neutral conformers of six alphabetic capital letters, A, B, C, D, E, and X, in Figure 4.1a(b) are quoted from report by Snoek **et al.**, in which their assignment was carried out based on a combination of UV-UV and IR-UV double-resonance

spectroscopy of jet-cooled L-phenylalanine and coupled with *ab* initio computations.[5] The nine Roman numerals I, II, ..., VIII, and IX are numbered in order of relative energies calculated using the MP2/6-311G** level of theory.

Neutral conformers IV, V, and VIII, which were predicted by the B3LYP results, have not been experimentally identified. The M05-2X results indicate that neutral conformer V belongs to subgroup II while cationic conformer V belongs to subgroup I.

This is totally different from the results obtained by B3LYP: both the neutral and the cationic conformers of V were assigned to subgroup II. Such a hybrid type of conformer that has different characters between the neutral and caion is one of the newly found results using the new DFT functional that takes into account noncovalent interactions as shown in Figure 4.1(b).

First, the conformer IV was assigned as observed structure E in ref 1, but IX was assigned later because of the obvious evidence in photo ionization energy (PIE) curve, shown by Lee **et al.**[7] According to ref 7, observed conformer E should belong to subgroup II based on the value of ionization energy whereas conformer IV belong to subgroup I, and instead conformer IX belong to subgroup II from the M05-2X results.

In previous papers on geometrical structures obtained by using the B3LYP method there was no clear difference in ionization energies or vibrational frequencies in the neutral ground state between V and VII conformers.[5,8,9]

Furthermore, there were no deviations of cationic structures, that is, the two conformers commonly showed elongation of the C_α–C_β bond and planarity of C_α–NH_2. (see Table 4.1) There has been heated debate about whether observed conformer A was to be assigned to conformer V or VII.[2,5]

Experimentally, conformer V was assigned to conformer A by Snoek **et al.**[5] Later, it was rationalized from UV rotational band contour analysis by ref 2 that conformer A corresponds not to conformer V but to conformer VII.

The results obtained by using the M05-2X method, in contrast, show no elongation of the C_α–C_β bond and a nonplanar, pyramidal structure for C_α–NH_2 on conformer V in Table 4.1.

Table 4.2 shows relative energies including zero-point energy correction (ZPE_{corr}) of the conformers, which were calculated by both methods. The original Roman numerals were used to avoid confusion even though their energy ordering is different from that of M05-2X in Table 4.2.

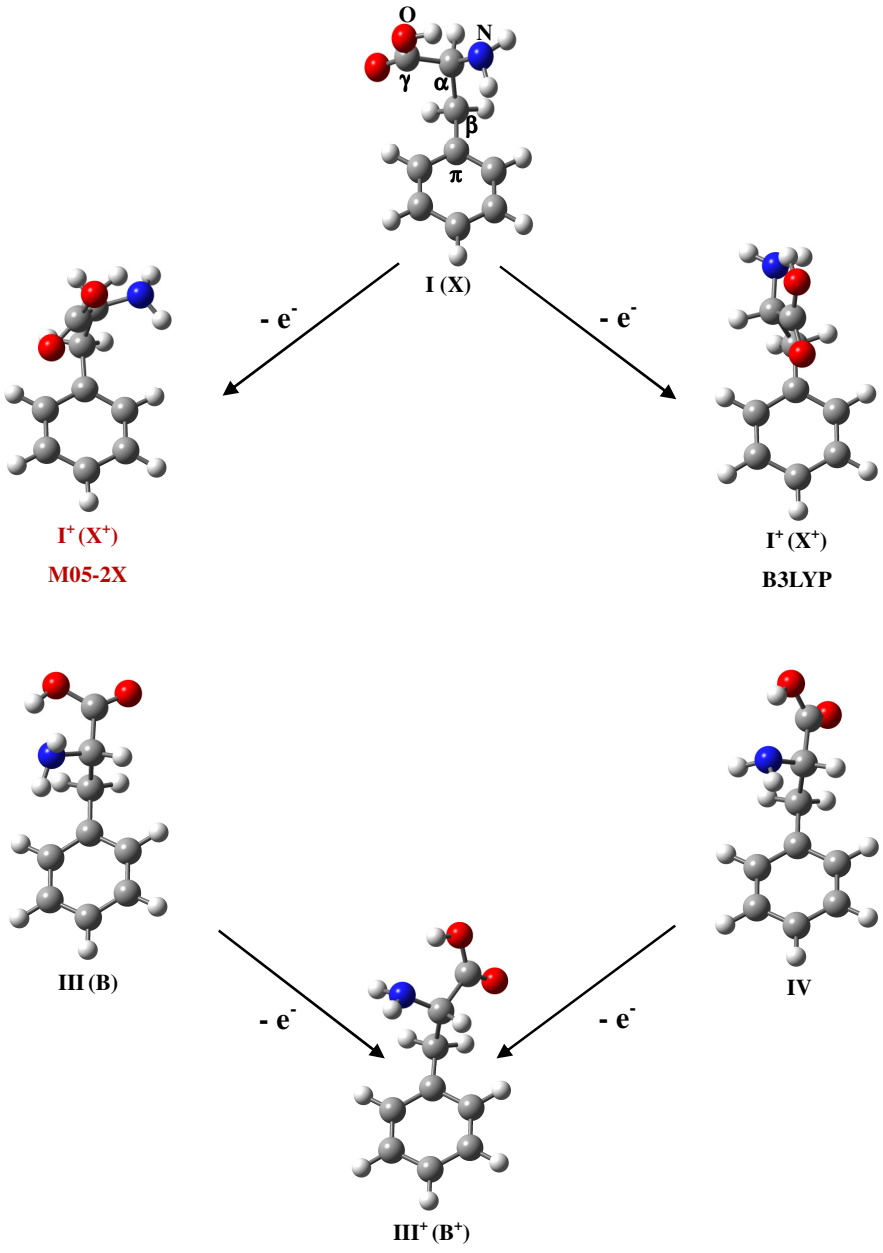


Figure 4.1(a) Three stable structures of conformers belonging to subgroup I in neutrals (cations) for L-phenylalanine.

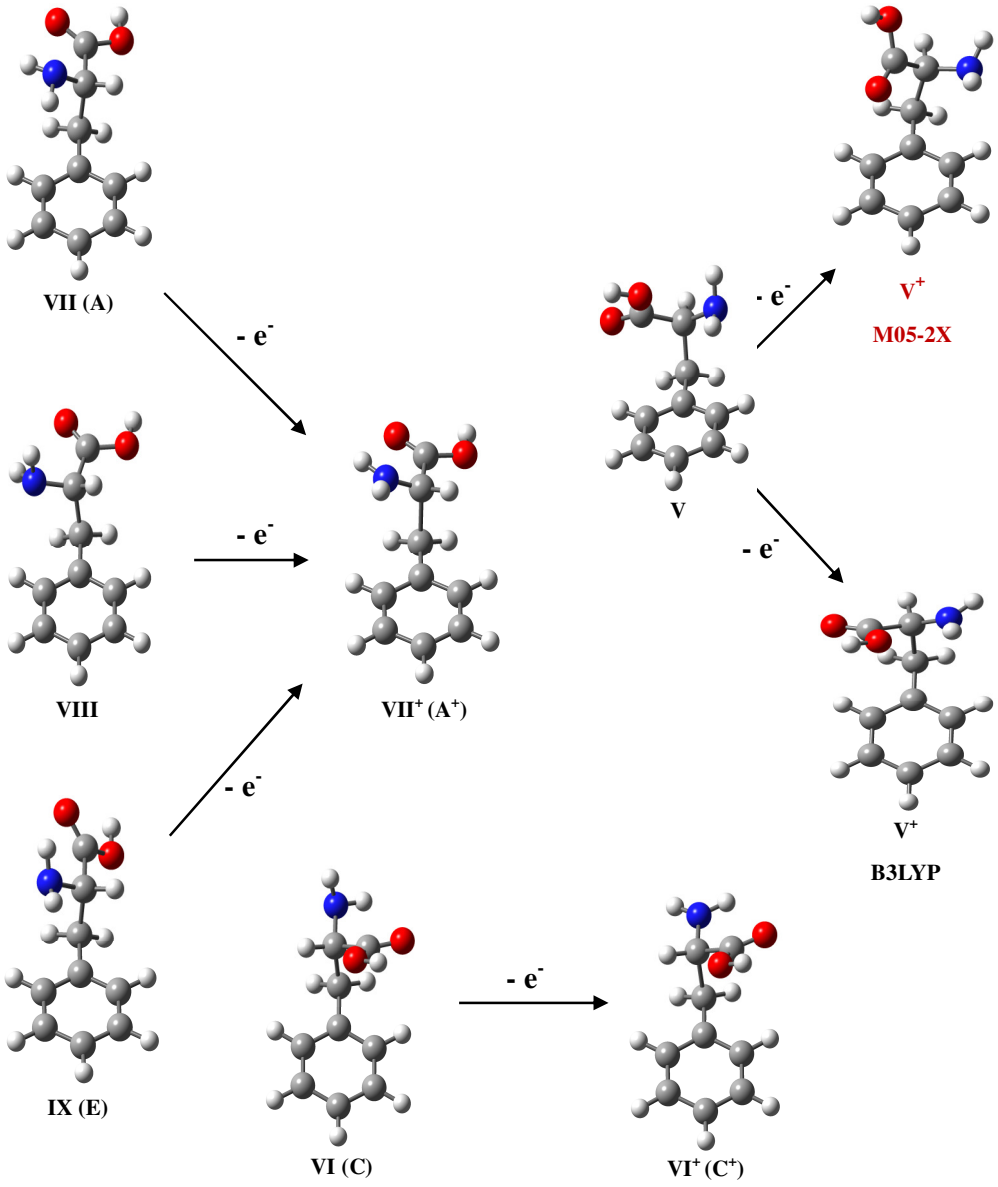


Figure 4.1(b) Six stable structures of conformers belonging to subgroup II in neutrals (cations) for L-phenylalanine.

Table 4.1. Calculated molecular constants for nine conformers of L-phenylalanine in neutrals and those for cations in parentheses.

(a) M05-2X, neutral (cation)

Sub.	Experiment Assignment	Conf.	Dihedral angle(°)		Bond length(Å)		
			C_π-C_β-C_α-N	C_α–NH_2	C_α–C_β	C_α–N	C_β– C_π
I	X	I	52(**88**)	123(131)	1.54(1.57)	1.46(1.44)	1.51(1.49)
I	B	III	-63(-86)	122(137)	1.54(1.59)	1.46(1.42)	1.51(1.46)
I		IV	-60(-86)	121(137)	1.53(1.59)	1.46(1.42)	1.51(1.46)
II (I)		**V**	62(64)	**120(127)**	**1.54(1.56)**	1.45(1.44)	1.51(1.48)
II	A	VII	-60(-69)	122(176)	1.55(1.77)	1.45(1.36)	1.50(1.45)
II	C	VI	178(171)	120(168)	1.54(1.79)	1.45(1.36)	1.51(1.44)
II	D	II	62(73)	119(178)	1.54(1.78)	1.45(1.36)	1.51(1.44)
II	E	IX	-62(-69)	123(176)	1.54(1.77)	1.45(1.36)	1.51(1.44)
II		VIII	-73(-69)	120(176)	1.54(1.77)	1.45(1.36)	1.51(1.45)

(b) B3LYP, neutral (cation)

Sub.	Experiment Assignment	Conf.	Dihedral angle(°)		Bond length(Å)		
			C_π-C_β-C_α-N	C_α–NH_2	C_α–C_β	C_α–N	C_β– C_π
I	X	I	52(**172**)	123(133)	1.55(1.56)	1.47(1.44)	1.51(1.50)
I	B	III	-64(-84)	122(143)	1.55(1.58)	1.47(1.42)	1.51(1.48)
I		IV	-64(-84)	122(143)	1.54(1.58)	1.47(1.42)	1.51(1.48)
II		**V**	61(72)	**121(179)**	**1.55(1.70)**	1.45(1.39)	1.51(1.47)
II	A	VII	-61(-72)	122(176)	1.57(1.70)	1.45(1.39)	1.51(1.47)
II	C	VI	-175(177)	121(165)	1.54(1.70)	1.46(1.39)	1.51(1.47)
II	D	II	62(73)	120(178)	1.55(1.70)	1.45(1.39)	1.51(1.46)
II	E	IX	-63(-73)	123(176)	1.55(1.70)	1.46(1.39)	1.51(1.46)
II		VIII	-73(-72)	120(176)	1.56(1.70)	1.45(1.39)	1.51(1.47)

Table 4.2. Relative energies in kcal/mol of the nine stable conformers in the neutral (cationic) ground state of L-phenylalanine.

Sub.	Experimental Assignment	Conf.	Neutral		Cation	
			M05-2X	B3LYP	M05-2X	B3LYP
I	X	I	0.00	0.11	6.29	1.68
I	B	III	0.62	0.00	10.03	7.29
I		IV	0.79	0.17	10.03	7.29
II		V	1.15	1.64	1.46	0.96
II	A	VII	1.50	1.00	0.71	0.00
II	C	VI	1.56	1.08	0.17	0.04
II	D	II	0.47	0.87	0.00	0.22
II	E	IX	1.71	1.22	0.71	0.00
II		VIII	2.06	1.09	0.71	0.00

(2) L-tyrosine

For L-tyrosine, eight band origins (4–7, 14–17) were observed by fluorescence-detected UV-UV and IR-UV double-resonance spectroscopy.[19] Figure 4.2a(b) show only the eight valid neutral conformers of L-tyrosine and their corresponding cations among the 18 optimized structures obtained by M05-2X: (a) for subgroup I, (b) for subgroup II.

Particularly, the cationic structures belonging to subgroup II obtained by the popular B3LYP are different from those from M05-2X and shown in parentheses.

A comparison of M05-2X and B3LYP calculations is shown in Table 4.3, which lists the dihedral angles ($N–C_\alpha–C_\beta–C_{\pi1}$, $C_\alpha–NH_2$, $N–C_\alpha–C_\gamma–O_1$) and the bond lengths ($C_\alpha–C_\beta$, $C_\beta–C_{\pi1}$, $C_\alpha–N$, and $C_{\pi2}–O_2$).

The two functionals, M05-2X and B3LYP, yielded nearly the same results with regard to the neutral structure, but the radical cations of subgroup II show different conformational structures depending on the functional used. This is mainly due to the effect of long-range noncovalent interactions, which is properly accounted for in the M05-2X functional but not in the B3LYP results.[21–24]

It should be noted in Table 4.3 that L-tyrosine conformers of subgroup II show no drastic change in geometry upon ionization, such as $C_\alpha–C_\beta$ elongation and planarity of the $C_\alpha–NH_2$ as seen in the subgroup II conformers of L-phenylalanine.[7,16] The neutral species shown in Figure 4.2a(b) are identical to

those reported by Ebata and co-workers.[19] The subgroup classification for L-tyrosine is based on the directionality of the intramolecular hydrogen bonding, as in the case of L-phenylalanine: $-COOH \rightarrow -NH_2$ for subgroup I and $-NH_2 \rightarrow -OCOH$ for subgroup II.

There are two different, nearly isoenergic, orientations of the phenolic –OH group, which are denoted by subscript r (right) or l (left). Roman numerals (I, II, III, and VII) were adopted from L-phenylalanine labels[16] because of the similarity between L-tyrosine and L-phenylalanine, with the –OH group in L-tyrosine being the major difference.

Ebata and co-workers observed eight origin bands (4, 5, 6, 7, 14, 15, 16, and 17) in their fluorescence excitation spectra[19] and proposed that pairs of 4 and 6, 5 and 7, and 16, and 17 correspond to two rotational isomers arising from the different orientations of phenolic –OH.

Table 4.4 shows relative energies of L-tyrosine conformers in neutrals (cations), which were calculated by using M05-2X and conventional B3LYP methods. Here, zero point energy corrections (ZPE_{corr}) were taken into account. According to the results in Table 4.4, these sets of rotational conformers are no different energetically.

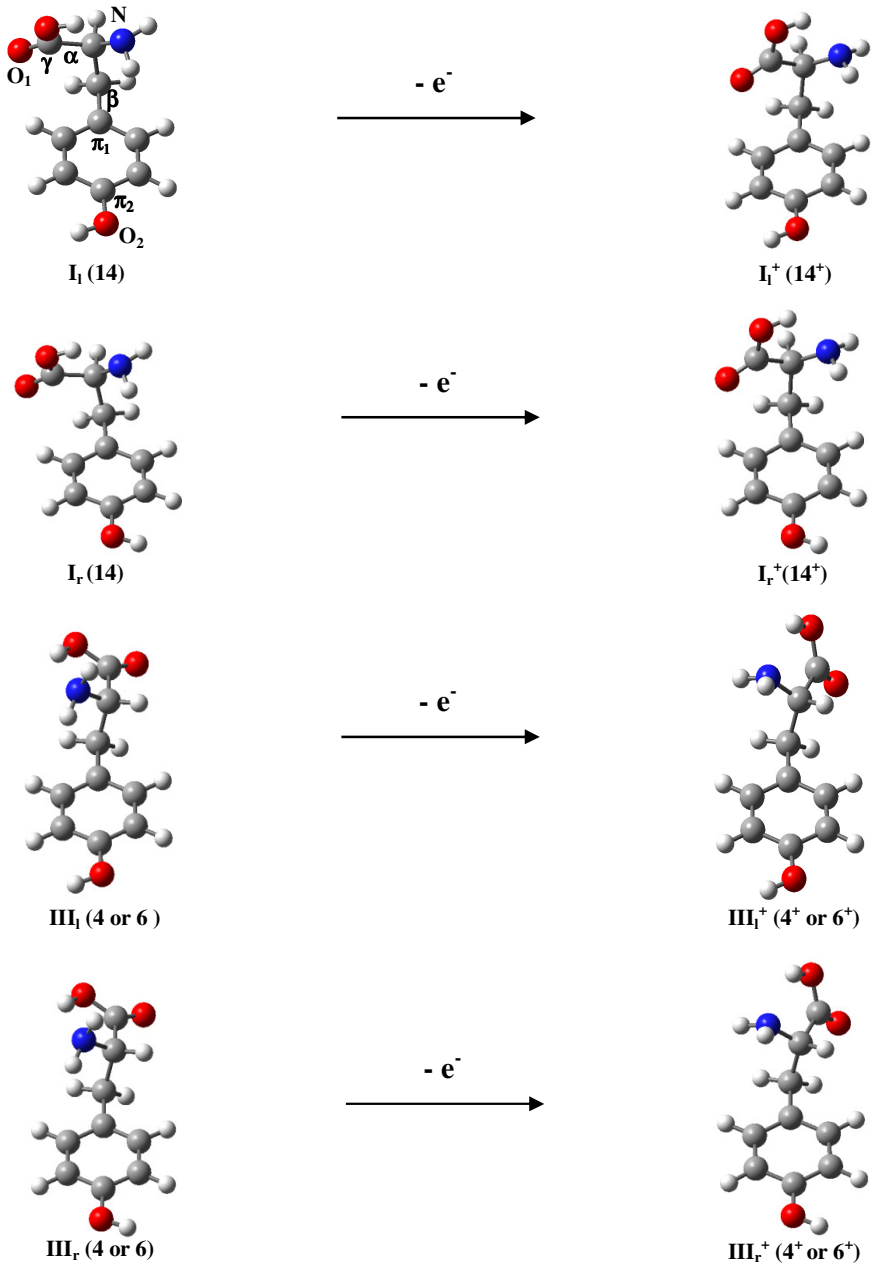


Figure 4.2(a) The four stable structures of representative L-tyrosine conformers belonging to subgroup I in neutrals (cations).

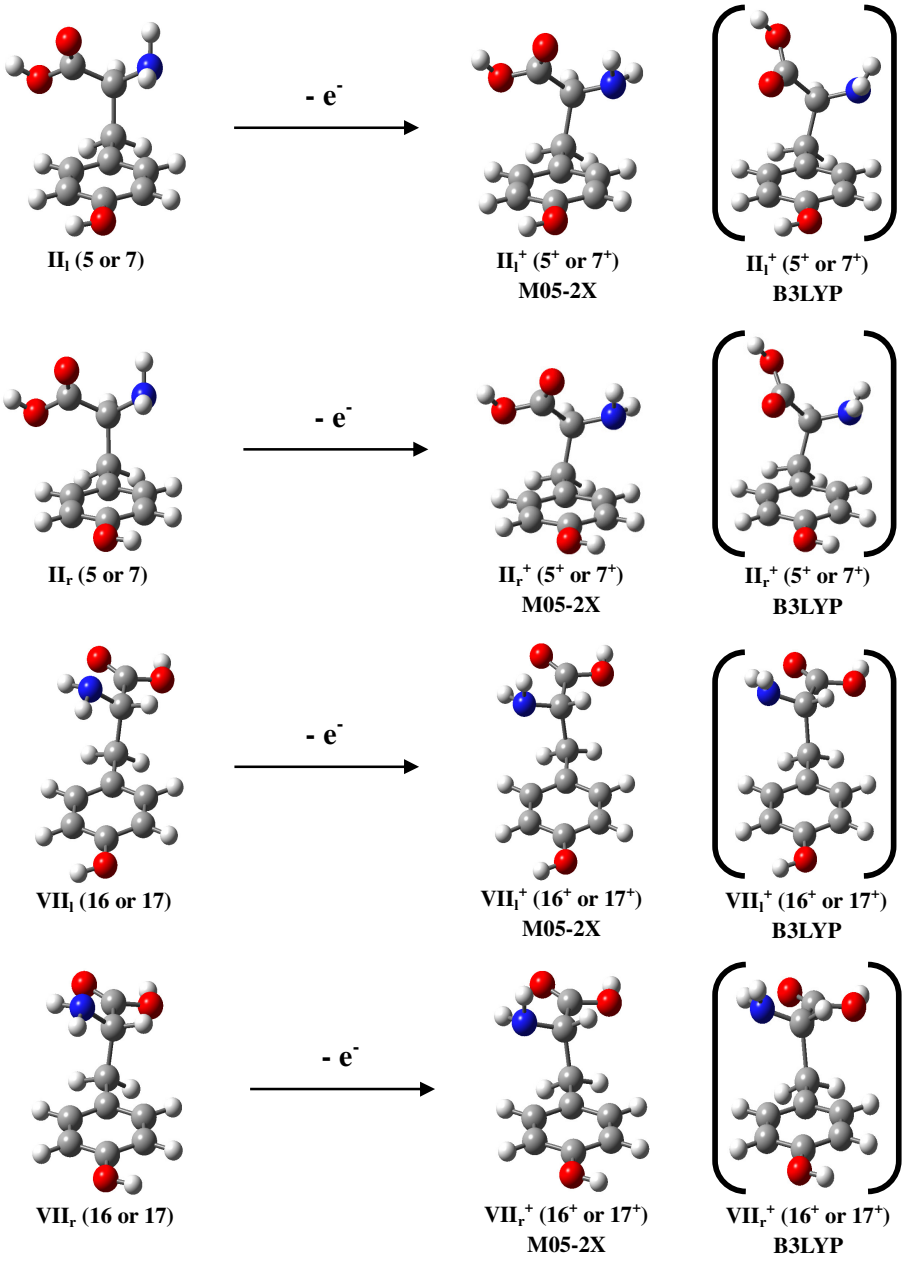


Figure 4.2(b) The four stable structures of representative L-tyrosine conformers belonging to subgroup II in neutrals (cations).

Table 4.3. Molecular constants for eight L-tyrosine conformers in neutrals and those for cations in parentheses.

Sub.	Origin Bands	Conf.	Dihedral angle (°)			Bond length (Å)			
			$C_{\pi 1}$-C_β-C_α-N	C_α-NH_2	O_1-C_γ-C_α-N	C_α-C_β	C_β-$C_{\pi 1}$	C_α-N	$C_{\pi 2}$-O_2
(a) M05-2X, neutral (cation)									
I	14	I_l	52 (70)	123 (132)	-169 (-150)	1.54 (1.56)	1.51 (1.48)	1.46 (1.44)	1.36 (1.31)
I	14	I_r	51 (68)	123 (132)	-169 (-151)	1.54 (1.56)	1.51 (1.48)	1.46 (1.44)	1.36 (1.31)
I	4 or 6	III_l	-63 (-80)	122 (132)	163 (-161)	1.54 (1.56)	1.51 (1.47)	1.46 (1.44)	1.36 (1.31)
I	4 or 6	III_r	-63 (-82)	122 (132)	163 (-162)	1.54 (1.56)	1.51 (1.47)	1.46 (1.44)	1.36 (1.31)
II	5 or 7	**II_l**	62 (66)	119 (125)	-1 **(0)**	1.54 (1.57)	1.51 (1.48)	1.45 (1.44)	1.36 (1.31)
II	5 or 7	**II_r**	62 (67)	119 (124)	0 **(0)**	1.54 (1.57)	1.51 (1.48)	1.45 (1.44)	1.36 (1.31)
II	16 or 17	**VII_l**	-60 (-63)	122 **(120)**	-30 (-12)	1.55 **(1.56)**	1.50 (1.48)	1.45 (1.45)	1.36 (1.31)
II	16 or 17	**VII_r**	-60 (-65)	122 **(121)**	-30 (-12)	1.55 **(1.56)**	1.50 (1.48)	1.45 (1.44)	1.36 (1.31)
(b)B3LYP, neutral (cation)									
I	14	I_l	53 (75)	123 (135)	-170 (-150)	1.55 (1.59)	1.51 (1.49)	1.47 (1.44)	1.37 (1.32)
I	14	I_r	52 (69)	123 (135)	-170 (-154)	1.55 (1.59)	1.51 (1.48)	1.47 (1.44)	1.37 (1.32)
I	4 or 6	III_l	-64 (-84)	122 (136)	165 (-162)	1.55 (1.58)	1.51 (1.48)	1.47 (1.43)	1.37 (1.32)
I	4 or 6	III_r	-64 (-85)	122 (136)	165 (-162)	1.55 (1.58)	1.51 (1.48)	1.47 (1.43)	1.37 (1.32)
II	5 or 7	**II_l**	62 (62)	120 (131)	-4 **(-86)**	1.55 (1.57)	1.51 (1.49)	1.45 (1.44)	1.37 (1.32)
II	5 or 7	**II_r**	62 (63)	120 (130)	-4 **(-81)**	1.55 (1.57)	1.51 (1.49)	1.45 (1.44)	1.37 (1.32)
II	16 or 17	**VII_l**	-61 (-73)	122 **(140)**	-32 (-17)	1.57 **(1.64)**	1.51 (1.47)	1.45 (1.41)	1.37 (1.32)
II	16 or 17	**VII_r**	-61 (-72)	122 **(143)**	-32 (-18)	1.57 **(1.64)**	1.51 (1.47)	1.45 (1.41)	1.37 (1.33)

Table 4.4. Relative energies (in kcal/mol) of eight stable L-tyrosine conformers in the neutral (cationic) gound state

Sub.	Origin Bands	Conf.	Neutral		Cation	
			M05-2X	B3LYP	M052X	B3LYP
I	14	I_l	0.00 (1)	0.00 (1)	4.60	4.69
I	14	I_r	0.32 (2)	0.35 (4)	5.53	5.51
I	4 or 6	III_l	1.05 (6)	0.13 (3)	8.70	5.56
I	4 or 6	III_r	0.82 (5)	0.11 (2)	8.45	5.37
II	5 or 7	II_l	0.56 (3)	0.91 (8)	0.18	0.00
II	5 or 7	II_r	0.66 (4)	0.90 (7)	0.00	0.38
II	16 or 17	VII_l	1.86 (13)	1.15 (9)	0.98	0.18
II	16 or 17	VII_r	1.65 (11)	1.15 (10)	1.30	0.04

(3) L-tryptophan

For L-tryptophan, six conformers (A–F) have been experimentally identified.[20] Figure 4.3 shows the geometrical structures of L-tryptophan for the most six valid neutral conformers and five cationic ones in their ground states. The structures were calculated by using M05-2X optimization with the 6-311+G* basis set and they were tentatively classified into two subgroups, I (I, III) and II (II, V), as for L-phenylalanine.[16]

It has already been qualitatively understood that each subgroup for L-phenylalanine has two different types of intramolecular hydrogen bonding.[5]

The two functionals, M05-2X and B3LYP, gave almost the same results regarding geometrical structures for L-tryptophan in neutral and cation. The neutral conformers of six alphabetic capital letters, A, B, C, D, E, and F in Figure 4.3 were quoted from ref 20 in which their assignment was carried out using mass-selected resonant two photon ionization (MS-R2PI) spectra, UV–UV hole-burn spectra, and IR–UV ion dip spectra together with predictions by *ab* initio MO calculations employing B3LYP.

The four Roman numerals, I, II, III, and V, were adopted from ref 16 on L-phenylalanine for similarity of the backbone, that is, orientation of backbone (alanine) in L-tryptophan and L-phenylalanine are almost the same. Subscripts a and b represent different orientations of the residue (indole ring) because of asymmetry in L-tryptophan.

In Figure 4.3, M05-2X results show that cationic structures of conformer I_a belonging to subgroup I are similar to that of L-phenylalanine, in other words, geometry of the backbone (alanine) in L-tryptophan is almost same as that of L-phenylalanine, while conformer III_a in radical cations is quite different from that of L-phenylalanine.

Table 4.5 shows the calculated molecular constants for six conformers of L-tryptophan in the neutral and radical cationic ground states. Both the M05-2X and B3LYP results have the similar tendency in their molecular constants.

Furthermore, there exist no significant differences in the molecular constants between subgroups I and II. Cationic L-tryptophan conformers belonging to subgroup II have neither C_α–C_β elongation nor planarity of C_α–NH_2, which is different from cationic conformers belonging to subgroup II of L-phenylalanine.

Table 4.6 shows relative energies of L-tryptophan conformers in neutrals (cations), which were calculated by using M05-2X and conventional DFT (B3LYP) method including zero point energy corrections (ZPE_{corr}). The M05-2X functional gave more reasonable candidates corresponding to observed conformers in L-tryptophan than the B3LYP based on order of relative energies.

As mentioned above, there exist two types of hydrogen bonding in L-tryptophan based on their structures and frequency analysis and then conformer I_a and III_a, which belong to subgroup I and which have intramolecular hydrogen bonding (–COOH → $–NH_2$), should be candidates for observed conformers A and F.

Conformer II_a, II_b, V_a, and V_b, which belong to subgroup II and which have free –OH, should correspond to conformers E, C, B, and D, which were observed by Simons and co-workers in ref 20. The results from newly employed M05-2X for L-tryptophan in neutral provides the validity to show good agreement with assignments by Simon and co-workers.

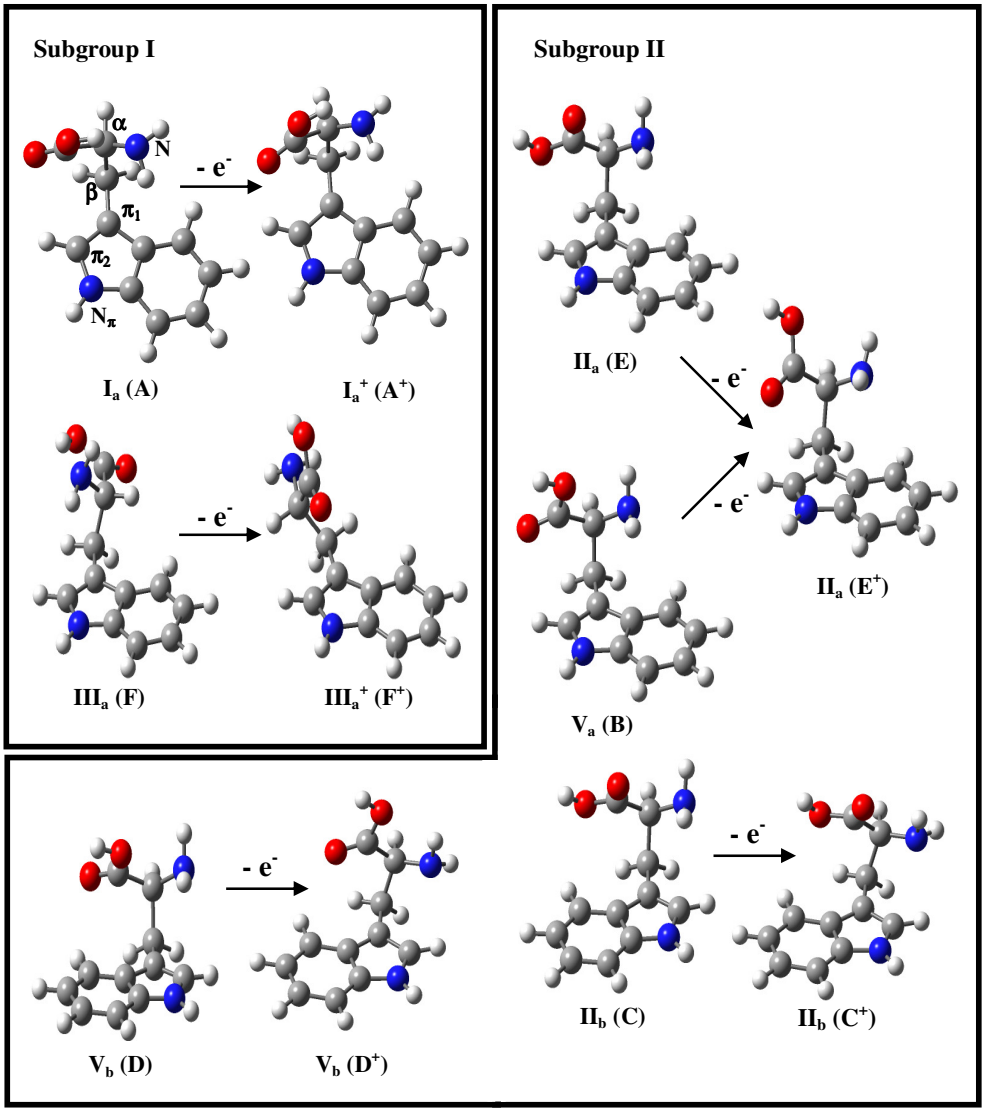


Figure 4.3. The six stable structures of representative L-tryptophan conformers in neutrals and cations.

Table 4.5. Molecular constants for six L-tryptophan conformers in neutrals and those for cations in parentheses.

Sub.	Experiment Assignment	Conf.	Dihedral angle (°)		Bond length (Å)			
			$C_{\pi1}$-C_β-C_α-N	C_α-NH_2	C_α-C_β	$C_{\pi1}$-$C_{\pi2}$	C_α-N	$C_{\pi2}$-N_π
(a) M05-2X, neutral (cation)								
I	A	I_a	55 (68)	122 (130)	1.53 (1.56)	1.37 (1.44)	1.46 (1.45)	1.37 (1.32)
I	F	III_a	-62 (172)	121 (128)	1.54 (1.54)	1.37 (1.43)	1.46 (1.45)	1.38 (1.32)
II	E	II_a	65 (58)	119 (123)	1.54 (1.54)	1.37 (1.43)	1.45 (1.45)	1.38 (1.32)
II	C	II_b	63 (63)	119 (122)	1.54 (1.55)	1.37 (1.43)	1.45 (1.45)	1.38 (1.32)
II	B	V_a	63 (58)	121 (123)	1.54 (1.54)	1.37 (1.43)	1.45 (1.45)	1.38 (1.32)
II	D	V_b	60 (59)	120 (120)	1.54 (1.54)	1.37 (1.43)	1.45 (1.46)	1.38 (1.32)
(b) B3LYP, neutral (cation)								
I	A	I_a	55 (72)	123 (131)	1.55 (1.57)	1.37 (1.43)	1.47 (1.45)	1.38 (1.33)
I	F	III_a	-62 (171)	121 (129)	1.55 (1.55)	1.37 (1.43)	1.47 (1.45)	1.38 (1.33)
II	E	II_a	65 (58)	120 (126)	1.55 (1.56)	1.37 (1.43)	1.45 (1.45)	1.38 (1.33)
II	C	II_b	64 (63)	119 (122)	1.55 (1.56)	1.37 (1.43)	1.45 (1.46)	1.38 (1.33)
II	B	V_a	64 (58)	121 (126)	1.55 (1.56)	1.37 (1.43)	1.46 (1.45)	1.38 (1.33)
II	D	V_b	60 (59)	121 (120)	1.55 (1.55)	1.37 (1.43)	1.45 (1.47)	1.38 (1.33)

Table 4.6. Relative energies (in kcal/mol) of six stable L-tryptophan conformers in the neutral (cationic) ground state.

Sub.	Experiment Assignments	Conf.	Neutral		Cation	
			M05-2X	B3LYP	M052X	B3LYP
I	A	I_a	0.00 (1)	0.00 (1)	4.14	4.12
I	F	III_a	1.39 (3)	0.13 (2)	4.36	2.55
II	E	II_a	1.32 (2)	0.91 (7)	0.00	0.00
II	C	II_b	1.66 (4)	0.90 (10)	0.26	0.98
II	B	V_a	1.83 (5)	1.15 (13)	0.00	0.00
II	D	V_b	2.11 (9)	1.15 (14)	0.23	0.94

4.3.2. Correspondence between the observed and computed conformers for L-phenylalanine

For L-phenylalanine, it has remained to be determined why conformers IV, V, and VIII, which were predicted by B3LYP calculations, have not been identified experimentally.

Conformer IV belonging to subgroup I has a relative energy (ΔE) of 0.17 kcal/mol (B3LYP) as shown in Table 4.2 and an activation energy E_a of 0.04 kcal/mol (see Figure 4.4(a)) in its neutral state, which is efficient for IV to be produced in a chamber under experimental conditions at temperatures of 170–190 °C (kT = 0.88 – 0.92 kcal/mol).

Conformer VIII belonging to subgroup II has also not been assigned experimentally, though B3LYP value of ΔE = 1.09 kcal/mol shown in Table 4.2 indicates the possibility of its observation. The M05-2X value of ΔE = 0.79 kcal/mol for conformer IV is very close to that of conformer III (B), ΔE = 0.62 kcal/mol. Here, entropy effects were neglected since the change in dihedral angle C_γ–C_α–C_β–C_π, which is associated with the backbone and residue rotation, is very small between the two conformers.

Figure 4.4a(b) show the results of intrinsic reacton coordinate (IRC) profiles for two conformerizations: III(B) – IV conformerization and VII(A) – VIII – IX(E) conformerization, respectively. Reaction paths through corresponding transition

state (TS) are determined by IRC. Here, zero-point energy correction (ZPE_{corr}) was taken into account. In Figure 4.4(a), III(B)–IV conformerization, both the M05-2X and B3LYP results gave almost the same behaviors for IRC profiles with a transition state (TS) having one imaginary frequency: the E_a of IV is 0.11 kcal/mol for M05-2X and 0.04 kcal/mol for B3LYP; ΔE between III(B) and IV is 0.17 kcal/mol for both M05-2X and B3LYP.

For III(B)–IV conformerization, reaction coordinate is rotation of C_α–C_α keeping intramolecular hydrogen bonding (–COOH → $–NH_2$). It has been shown that complex molecules with multiple degrees of conformational flexibility need a barrier height of 600 cm^{-1} (1.72 kcal/mol) or higher to be trapped in jet expansion experiments.[17]

This shows that under ordinary experimental conditions for a supersonic jet expansion experiment it is highly possible for III (B) and IV to be detected as the same conformer. The population ratio between III(B) and IV is roughly estimated to be 1:0.21 using a vibrational temperature of T = 57 K, which was adopted from rotational band-counter analyses of conformers in a supersonic jet expansion experiment.[5]

The ratio obtained by assuming that conformers retain thermal equilibrium is approximate, and detailed discussions are needed on cooling and trapping processes in supersonic jet expansion experiments in order to obtain more reliable population ratios.[17,18]

Although it is difficult to find clear evidence of the coexistence of the two conformers from the corresponding band in UV spectra, the possibility of coexistence is indicated by comparison of experimental rotational band contours with theoretically predicted results.[2] Here, the experimental band contours cannot be fitted well to the theoretical ones within one conformer III(B) model.

For 4.4(b) VII(A)–VIII–IX(E) conformerization, rotation of C_α–NH_2 is reaction coordinate. Two results are totally different: the B3LYP cannot explain the fact that VIII have not been observed. On the other hand, its M05-2X calculation showed ΔE = 2.06 kcal/mol ≈ 700 cm^{-1}, indicating that the energy is too high for it to be observed. The population of conformer VIII, which was estimated in the same way as described above, is too small to be observed, while VII(A) and IX(E) can be observed.

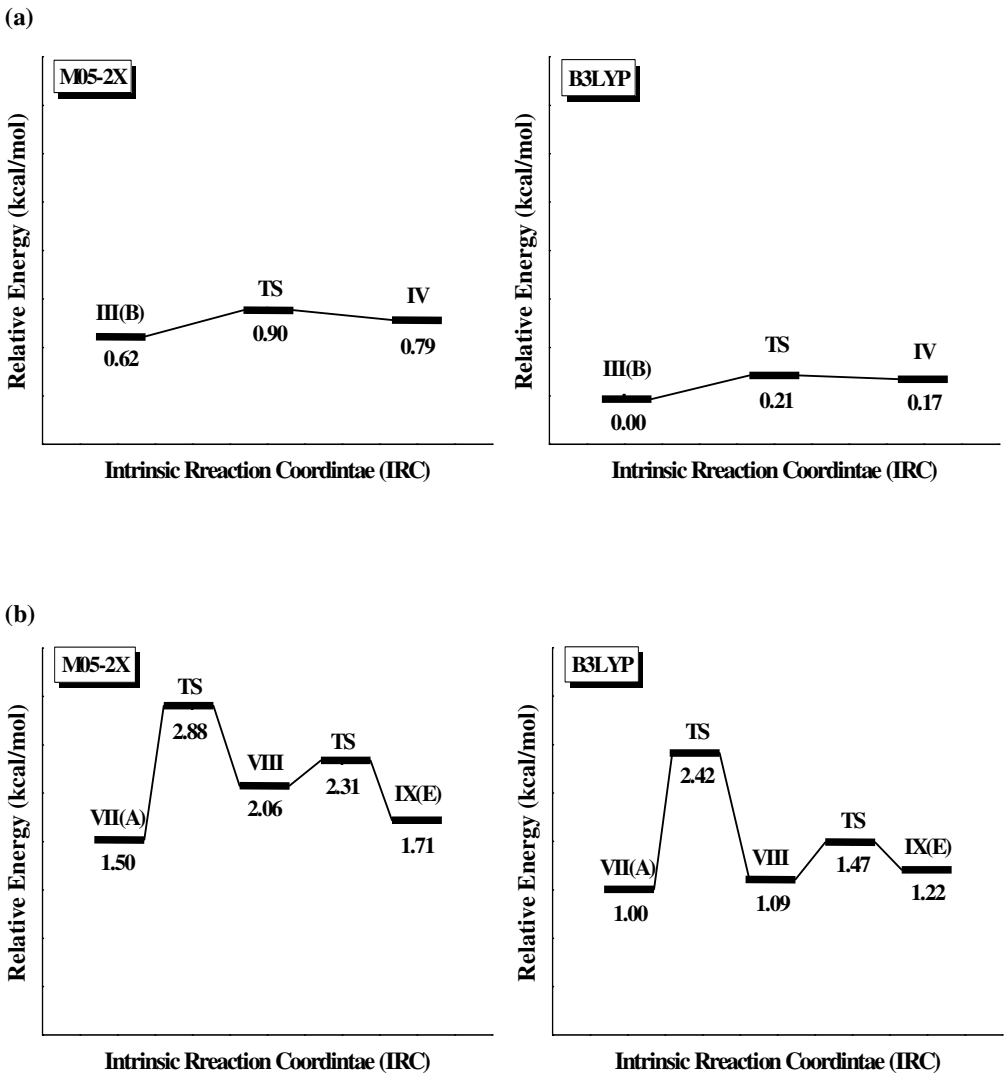


Figure 4.4. Intrinsic reaction coordinate (IRC) profiles on conformerizations of L-phenylalanine conformers in neutrals. (a) III(B)–IV conformerization from M05-2X and B3LYP; (b) VII(A)–VIII–IX(E) conformarization from M05-2X and B3LYP.

4.4. Conclusion

Firstly, B3LYP calculations were carried out in order to confirm the previous classification of the conformers into two subgroups (I and II) and then M05-2X calculations were conducted for both the neutral and the cationic conformers of all the aromatic amino acids. The two functionals, M05-2X and B3LYP, gave the same nine stable conformers in their neutral state.

However, quantitatively, their relative energies were different, the magnitudes of which are sufficient to identify the conformers: the M05-2X results can unambiguously explain the correspondence between the observed and the predicted conformers in terms of the relative energy and activation energy between the two conformers.

Especially, some cationic conformers were shown significantly different structure between two methods: for L-phenylalanine, conformer I and V in Figure 4.1a(b), for L-tyrosine, all conformers belonging to subgroup II in Figure 4.2(b). The difference results from absence of noncovalent intramolecular interaction in B3LYP.

By showing our M05-2X results, it is strongly emphasized to include long-range noncovalent effects for study of intramolecular hydrogen bonding systems like aromatic amino acids.

Here, it is newly suggested that Conformer V is of a hybrid character, belonging to subgroup I in its neutral form and belonging to subgroup II in its cation form. This character should be confirmed by experiments.

On the basis of the intrinsic reaction coordinate (IRC) results predicted by the new functional, it can be inferred why conformers IV, V, and VIII, which were predicted by B3LYP calculations, have not been identified experimentally.

4.5. References

[1] Zhao, Y; Truhlar, D. G. *J. Phys. Chem. A* 2004,*108,* 6908..

[2] Lee, Y.; Jung, J.; Kim, B.; Butz, P.; Snoek, L. C.; Kroemer, R. T.; Simons, J. P. *J. Phys. Chem. A* **2004**, *108,* 69.

[3] Lee, K. T.; Sung, J.; Lee, K. J.; Park, Y. D.; Kim, S. K. *Angew. Chem. Int. Ed.* **2002,** 41, 4114.

[4] Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321,* 49.

[5] Bouchoux, G.; Bourcier, S.; Blanc, V.; Desaphy, S. *J. Phy. Chem. B* **2009**, *113,* 5549.

[6] Lee, K. T.; Sung, J.; Lee, K. J.; Kim, S. K.; Park, Y. D. *Chem. Phys. Lett.* **2003**, 368, 263

[7] Lee, K. T.; Kim, H. M.; Han, K. Y.; Sung, J.; Lee, K. J.; Kim, S. K. *J. Am. Chem. Soc.* **2007**, *129*, 2588.

[8] Hashimoto, T.; Takasu, Y.; Yamada, Y.; Ebata, T. *Chem. Phys. Lett.* **2006**, *421,* 227.

[9] Huang, Z.; Yu, W.; Lin, Z. *Theochem-J. Mol. Struct.* **2006**, *758,* 195.

[10] Ebata, T.; Hashimoto, T.; Ito, T.; Inokuchi, Y.; Altunsu, F.; Brutschy, B.; Tarakeshwar, P. *Phys. Chem. Chem. Phys*. **2006**, *8,* 4783.

[11] Ebata, T.; Hashimoto, T.; Ito, T.; Inokuchi, Y.; Altunsu, F.; Brutschy, B.;

Tarakeshwar, P. *Phys. Chem. Chem. Phys*. **2006**, *8,* 4783.

[12] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[13] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[14] Gaussian 09, Revision A.1, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2009**.

[15] Huang, Z.; Yu, W.; Lin, Z. *Theochem-J. Mol. Struct.* **2006**, *758,* 195.

[16] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583.

[17] Godfrey, P. D.; Brown, R. D.; Rodgers, F. M. *J. Mol. Struct.* **1996**, 376, 65.

[18] Florio, G. M.; Christie, R. A.; Jordan, K. D.; Zwier, T. S. *J. Am. Chem. Soc.* **2002**, *125*, 10236.

[19] Inokuchi, Y.; Cobayashi, Y.; Ito, T.; Ebata, T. *J. Phys. Chem. A* **2007**, *111,* 3209.

[20] Snoek, L. C.; Kroemer, R. T.; Hockrige, M. R.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3,* 1819.

[21] Zhao, Y.; Truhlar, D. G. *Accounts Chem. Res.* **2008**, *41,* 157.

[22] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[23] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123,* 161103.

[24] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[25] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2011**, *115*, 9658.

Chapter 5.

A Tool Investigating

Conformation-Dependent Property:

Ionization Energy

5.1. Introduction

Ionization energy (IE) means the minimum quantity of energy required in order to remove one electron from an atom or molecule. This is the definition of adiabatic ionization energy. In the process of adiabatic ionization, molecular geometry should be changed.

Vertical ionization energy is defined as the energy gap between neutral and cation states under the nucleus frozen condition. It is concerned with Franck-Condon overlap and molecular geometry is fixed on the ionization. The definition of adiabatic and vertical ionization energies are illustrated in Figure 5.1.

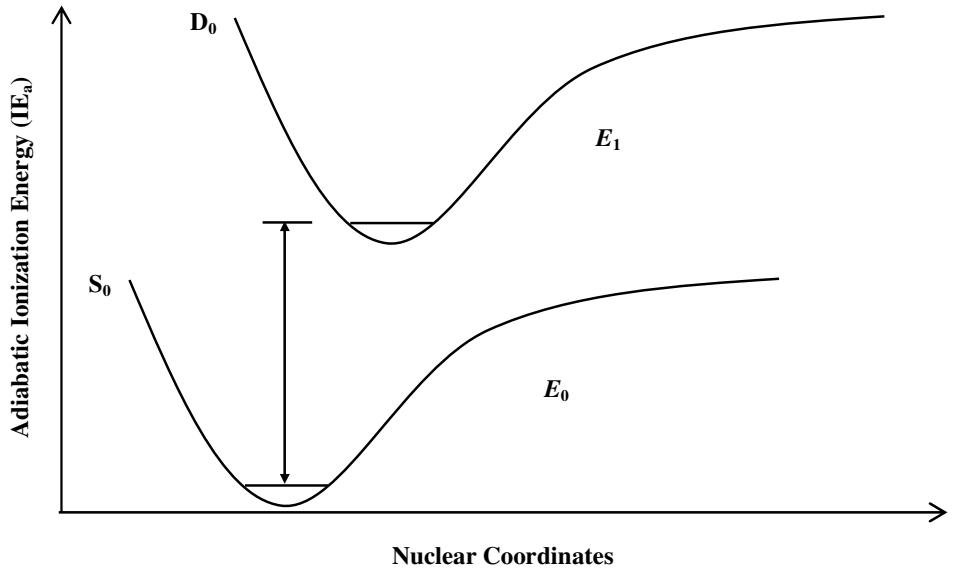
Photoionization, a kind of the ionization methods, is described to be easy understand in ref 1 and it utilizes electromagnetic radiation or photon in order to provide energy source to ionized neutral species.

Nowadays, a variety of spectroscopic techniques have been developed to ionize molecules and one of the most popular way is resonant two-photon ionization (R2PI) technique: the wavelength of the first photon is tuned to a resonant transition and the subsequent second photon brings the molecule above the ionization threshold.

By using this technique, the ion signal rise abruptly from zero, the onset of which can be regarded as the IE. This spectrum is called photoionization–efficiency (PIE) curve. The PIE curve of the result obtained for L-phenylalanine

shows that there is strong conformation-dependent correlation between the ionization energies of molecules and intramolecular hydrogen bonding interactions.[2]

In the next section, theoretical methods and procedures are briefly described. In section 5.3, vertical and adiabatic ionization energies of conformers for the three aromatic amino acids are calculated to provide further evidence to support the regrouping of subgroups I and II.



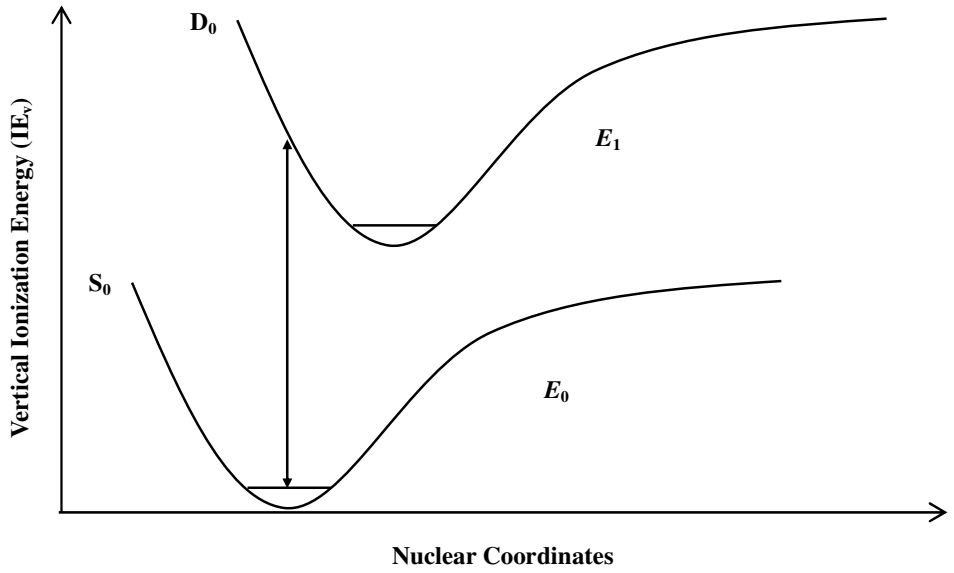


Figure 5.1. Ionization energy diagram.

5.2. Computational procedure

Firstly, geometry optimization is carried out for all neutral conformers of the three aromatic amino acids in ground state by setting the charge and the multiplicity to 0 and 1 at M05-2X/6-311+G* level of theory.

In the second place, it is evaluated by single point energy calculation for vertical ionization of each conformer at the optimized neutral in ground state by setting the charge and the spin multiplicity to +1 and 2 at M05-2X/6-311+G* level, respectively.

Whereas, geometry optimization for each cationic conformer of the three aromatic amino acids in their ground must be performed to evaluate the value of adiabatic ionization energy.

Here, all the calculations for ionization energy include zero-point energy corrections (ZPE_{corr}). The calculation at B3LYP/6-311+G* also is applied all calculation procedure for ionization energy with M05-2X method, simultaneously.

5.3. Results and discussion

5.3.1. L-phenylalanine

Figure 5.2a(b) shows calculated adiabatic (vertical) ionization energies of nine L-phenylalanine conformers. Both the adiabatic and vertical ionization energies calculated with M05-2X are higher than those calculated with B3LYP by ~0.5 eV for all of the conformers.

There have been no experimental reports on adiabatic ionization energies except for photoelectron measurements. This is because Franck-Condon factors for 0-0 transitions are negligibly small, which are due to large geometrical changes between the neutral and the cationic states, that is, elongation of the $C_\alpha-C_\beta$ bond and dihedral angle of $N-C_\alpha-C_\beta-C_\pi$ and pyramidal structure for $-NH_2$ on ionization as shown in Table 4.1.

Photoelectron spectroscopic values of toluene (8.82 eV)[5] and L-phenylalanine (8.50 eV)[6] are plotted in Figure 5.2(a). The predicted adiabatic ionization energies, I_a (M05-2X and B3LYP) for toluene are also plotted to estimate a possible error in the calculations. The M05-2X adiabatic ionization energies of the subgroup I conformers, I (X), III(B), and IV, are close to that of toluene, and the adiabatic ionization energies of the subgroup II conformers, VII(A), VI(C), II(D), and IX(E), are close to that of L-phenylalanine by photoelectron measurement, and

their energy differences are within 0.1eV. The adiabatic energies calculated B3LYP, on the other hand, are far from the experimental values by ~0.6 eV for the subgroup I conformers and ~0.4 eV for the subgroup II conformers.

Concerning vertical ionization energies, conformational-dependent ionization energies have been reported by *Lee* **et al.**[7] They measured photoionization efficiency (PIE) curves as a function of the wavelength of the ionization laser, while fixing the wavelength of the excitation laser at the S_1 band origin of each conformer in mass-resolved resonant two-photon ionization (R2PI) experiments.

The ion signal in PIE curve appears at ~8.8 eV for all the subgroup II conformers (VII(A), VI(C), II(D), and IX(E)) and at ~9.0 eV for I(X) and 9.15 eV for III(B) of subgroup I conformers. These correspond to the vertical ionization energies. The values thus estimated are indicated by the solid squares in Figure 5.2(b). Only one experimental value of 9.4 eV, which was obtained from a photoelectron spectrum[8] of a mixture of L-phenylalanine conformers, is given.

The estimated values for the vertical ionization energies of subgroup I conformers are marked by solid squares with an arrow, which represent lower bounds for the vertical ionization energies observed because of restriction in the wavelength ranges in the PIE measurement.

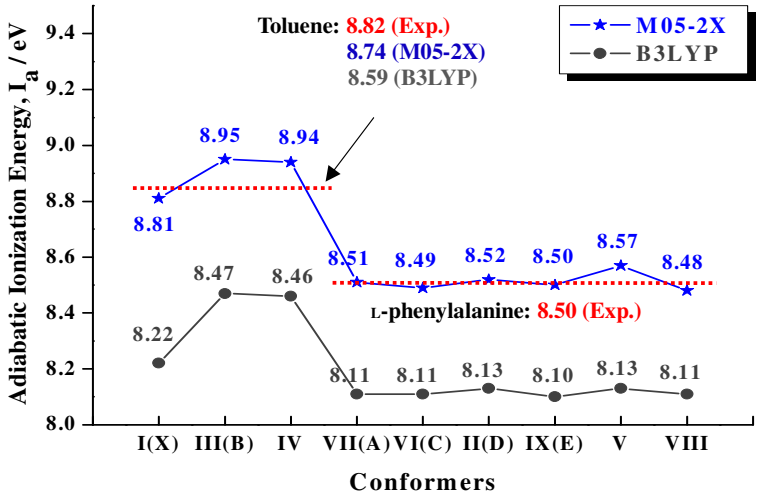
It is demonstrated that the M05-2X functional satisfactorily reproduces observed vertical ionization energies, being close to both the PIE and the

photoelectron experimental values, while the B3LYP functional could not quantitatively reproduce these experimental values.

It has already been pointed out that there is a tendency in ionization energy differences between subgroup I and II conformers, that is, the ionization energies of subgroup I conformers are higher than those of subgroup II.

It is important that ionization energy can be utilized as a tool to verify the conformation-dependent properties. As shown in Figure 5.2a(b) (M05-2X) and ref 7, one can understand there exists different type of structure between conformers of L-phenylalanine based on their tendencies of energy gap.

(a)



(b)

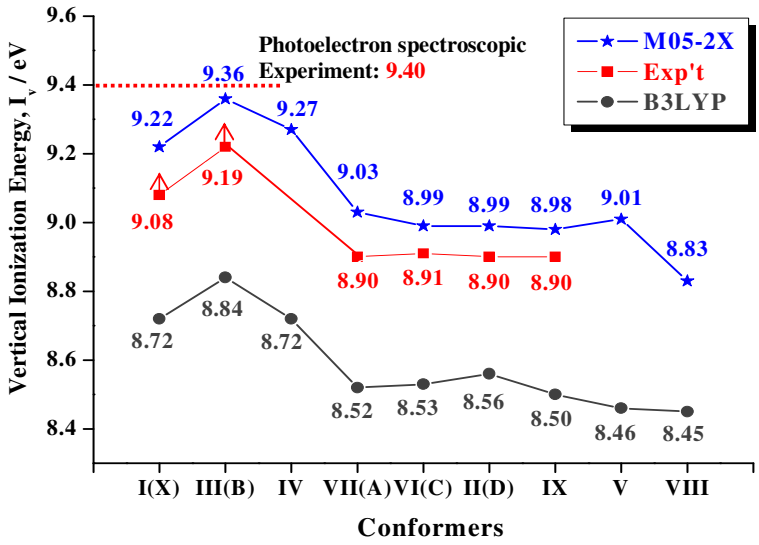


Figure 5.2. M05-2X and B3LYP results for (a) adiabatic and (b) vertical ionization energies for the nine most stable conformers of L-phenylalanine.

5.3.2. L-tyrosine and L-tryptophan

Figure 5.3 shows the calculated adiabatic ionization energies for L-tyrosine in the middle and those for L-tryptophan at the bottom. For comparison, in the upper part, adiabatic ionization energies for L-phenylalanine[3] are also shown. The adiabatic ionization energies calculated by M05-2X are higher than those calculated by B3LYP for all of the conformers by nearly 0.4 eV. Here, it can be seen that adiabatic ionization energies for all of the conformers of both L-tyrosine and L-tryptophan can be classified into two groups depending on subgroups I and II as those for L-phenylalanine: adiabatic ionization energies for the subgroup I conformers are higher than those for the subgroup II conformers. The energy difference between subgroup I and subgroup II becomes small, i.e., 0.29, 0.21, and 0.19 eV, as aromatic amino acid is replaced from L-phenylalanine to L-tyrosine and L-tryptophan.

Experimental values of adiabatic ionization energy are also shown in Figure 5.3. The adiabatic ionization energies of subgroup II (II and VII) conformers of L-tyrosine, which were calculated by M05-2X, reproduces those observed by photoelectron spectroscopy on L-tyrosine,[6] whereas for conformer I and III belonging to subgroup I, their adiabatic ionization energies reproduce photoelectron spectroscopic values of cresol (8.35 eV)[5] and L-tyrosine (8.00 eV)[6] are plotted in the middle of Figure 5.3.

Both the M05-2X and the B3LYP values of adiabatic ionization energy for cresol are also plotted to estimate a possible error in the calculations.

For L-tryptophan, the M05-2X adiabatic ionization energies of the subgroup I conformers, I_a (A) and III_a (F), are close to that of skatol at the bottom of Figure 5.3. The adiabatic ionization energies of the subgroup II conformers, II_a (E), II_b (C), V_a (B), and V_b (D), are close to that of L-tryptophan and their energy differences are within 0.10 eV.[10]

Alanine is the backbone common to all of the three aromatic amino acids, and the experimental values of their adiabatic ionization energies are 8.82, 8.88, 8.35, and 7.51 eV for toluene,[5] alanine,[6] cresol,[11] and skatol,[12] respectively.

The ionization energies predicted by M05-2X (B3LYP) with the 6-311+G* basis set are 8.74 (8.59), 8.12 (7.95), and 7.53 (7.31) eV for toluene, cresol, and skatol, and 9.36 (9.07), 8.94 (8.70) eV for alanine belonging to subgroups I and II, respectively. Photoelectron spectroscopic values of skatol (7.51 eV)[12] and L-tryptophan (7.30 eV)[10] are plotted in the bottom of Figure 5.3. Both the M05-2X and B3LYP values of adiabatic ionization energy for skatol are also plotted to estimate a possible error in their calculations.

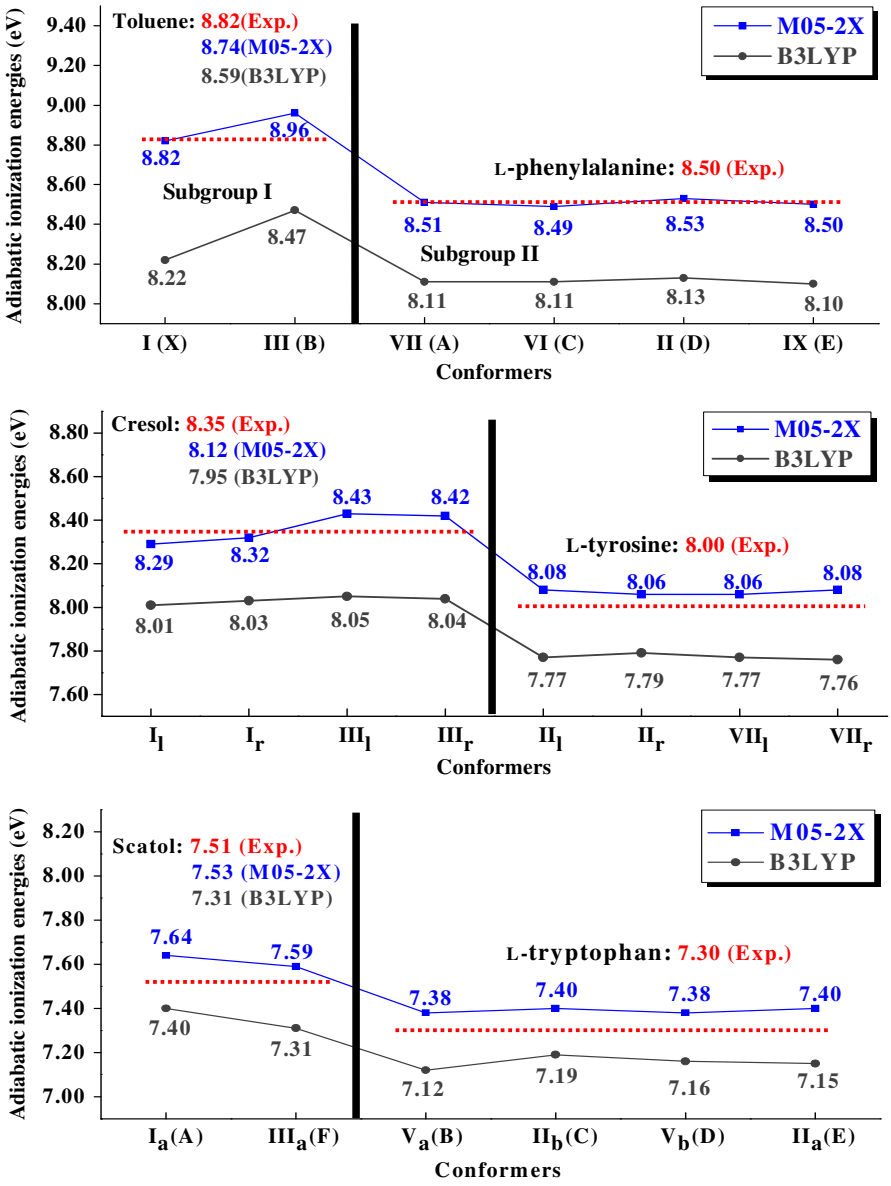


Figure 5.3. M05-2X and B3LYP results for adiabatic ionization energies of the six stable conformers in L-phenylalanine (upper), eight stable conformers in L-tyrosine (middle), and six stable conformers in L-tryptophan (lower).

Figure 5.4 shows a qualitative explanation for the calculated adiabatic ionization energies of each aromatic amino acid based on a simple molecular orbital picture.

Figure 5.4(a) shows the ionization of the three aromatic amino acids belonging to subgroup I. The tendency can be explained in terms of intramolecular hydrogen bonding between lone pair electrons of the amino group and hydrogen of the carboxyl group for subgroup I conformers. The n-MO of the amino group is stabilized as a result of hydrogen bonding, and interactions between the π-HOMOs of the phenyl group and the n-MO become negligibly small.

Figure 5.4(b) shows the ionization of the aromatic amino acids belonging to subgroup II. In this case, on the other hand, ionization takes place from the HOMO of aromatic amino acids.

In this section, I try to discuss the characteristic only shown in subgroup II of L-phenylalanine using a model in 5.4(b). Only subgroup II of L-phenylalanine, the energy levels of π-MOs and n-MO are close to each other as can be seen in Figure 5.4(b), and a pair of MOs is formed between a π-MO and an n-MO.

It results from their submoiety having almost same ionization energy: (toluene: 8.82 eV, alanine: 8.88eV) as mentioned above. This newly formed MOs are called hyperconjugative MOs.[9] Ionization from one of the hyperconjugative MOs with an antibonding character results in lowering of ionization energies of subgroup II conformers. C_α–C_β elongation of the subgroup II conformers on ionization can

qualitatively be explained in terms of electron dynamics associated with hyperconjugative MO.

The C_α–C_β elongation is mainly brought about by an electron release from the antibonding hyperconjugative MO: after ionization, both chromophores, phenyl and amino groups, have a positive partial charge with equal weights and attract electrons in the C_α–C_β bond. As a result, elongation occurs.

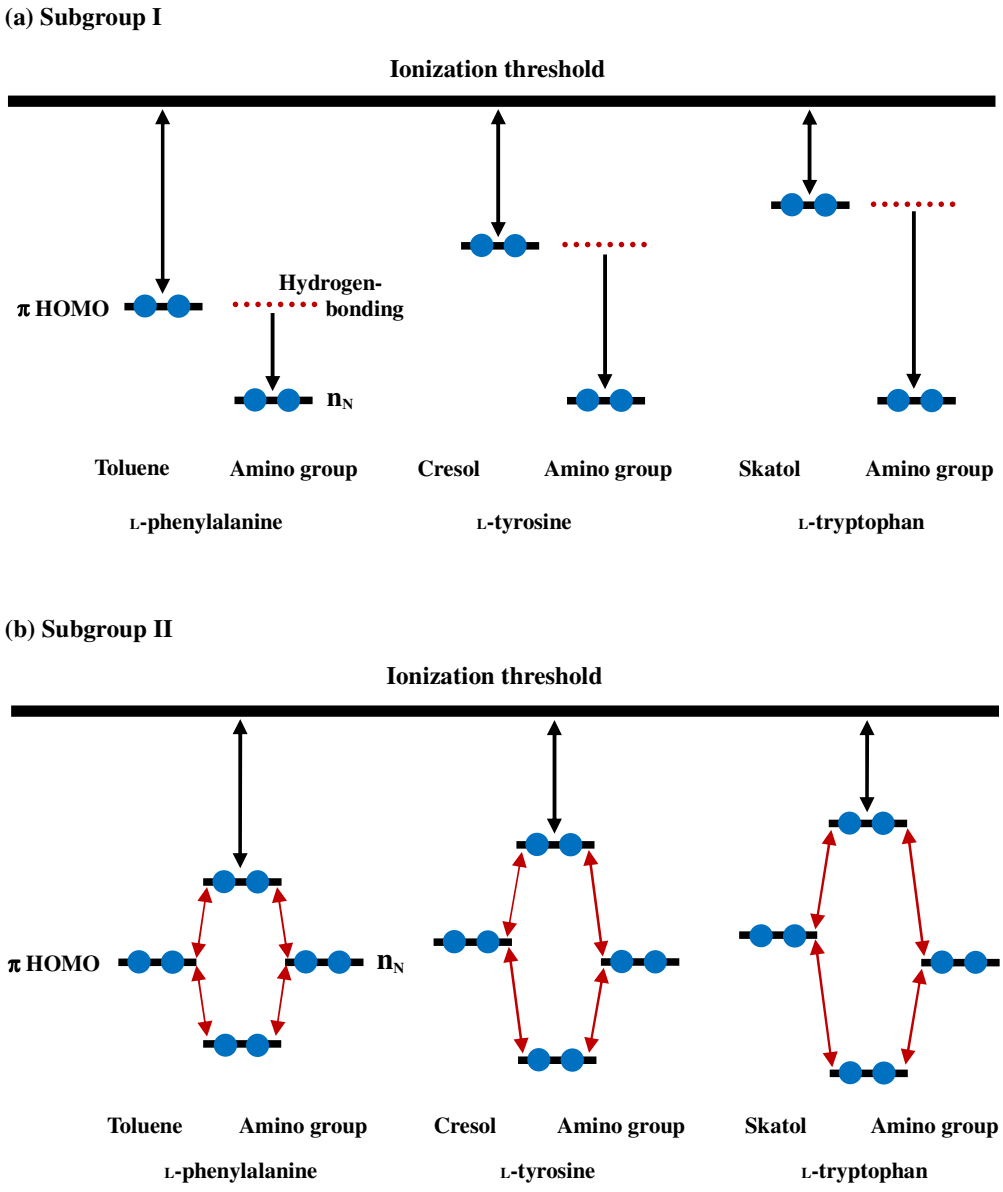


Figure 5.4. An MO picture showing the differences in ionization energies between conformers belonging to subgroup I and those belonging to subgroup II for the three aromatic amino acids

5.4. Conclusion

In this chapter, M05-2X calculations gave adiabatic (vertical) ionization energies in cationic conformers for all the three aromatic amino acids and several chromophores in order to explain the correlation between the conformational dependent properties of molecule and ionization energies, theoritically.

Based on the M05-2X results, unique character of subgroup II for L-phenylalanine could be describe by employing a simple MO model. Figure 5.4 can explain the reason why only L-phenylalanine conformers belonging to subgroup II against other aromatic amino acids (L-tyrosine and L-tryptophan) shows significant elongated bond length of the C_α–C_β and planarity of $-NH_2$ when they became cations.

The present results provide a reliable basis for analysis of ionization efficiency spectra and investigation of mechanisms of chemical reactions of the aromatic amino acids. Ionization-induced conformational change is one of the examples of chemical reactions.[13,14]

5.5. References

[1] Lee, K. T. Ph. D. Dissertation, Seoul National University, Seoul, 2002. pp.20-24.

[2] Lee, K. T.; Sung, J.; Lee, K. J.; Park, Y. D.; Kim, S. K. *Angew. Chem. Int. Ed.* **2002**, *41,* 4114.

[3] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583

[4] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2011**, *115*, 9658.

[5] Lu, Kuelh-Tzu; Eiden, G. C.; Weisshaar, J. C. *J. Phys. Chem.* **1992,** *96,* 9742.

[6] Campbell, S.; Beauchamp, J. L.; Rempe, M.; Lichtenberger, D. L. *Int. J. Mass Spectrom. Ion Processes* **1992,** *117*, 83.

[7] Lee, K. T.; Sung, J.; Lee, K. J.; Park, Y. D.; Kim, S. K. *Angew. Chem. Int. Ed.* **2002**, *41,* 4114.

[8] Cannington, P. H.; Ham, N. S. *J.Electron Spectrosc. Relat. Phenom.* **1983**, 32, 139.

[9] Sahnoun, R.; Fujimura, Y.; Kabuto, K.; Takeuchi, Y.; Noyori, R. *J. Org. Chem.* **2007**, *72,* 7923.

[10] Wilson, K. R.; Jimenez-Cruz, M., Nicolas, C.; Belau, L.; Leone, S. R.; Ahmed, M. *J. Phys. Chem. A* **2006**, *110,* 2106.

[11] Palmer, M. H.; Moyes, W.; Speirs, M.; Ridyard, J. N. A. *J. Mol. Struct.* **1979**, *52*, 293.

[12] Hager, J. W.; Wallace, S. C. *Anal. Chem.* **1988**, *60*, 5.

[13] Miyazaki, M.; Fujii, A.; Ebata, T.; Mikami, N. *Phys. Chem. CHem. Phys.* **2003**, *5*, 1137.

[14] Ishiuchi, S. –i.; Sakai, M.; Tsuchida, Y.; Takeda, A.; Kawashima, Y.; Fujii, M.; Dopfer, O.; Müller-Dethlefs, K. *Angew. Chem. Int. Ed.* **2005**, 44, 6149.

Chapter 6.

Hyperconjugation Effects on Charge Distribution in the Cationic Aromatic Amino Acids

6.1. Introduction

Weak interaction dominated by dispersion and intramoleular hydrogen bonding are important for van der Waals molecules and biological systems. The molecular shape and conformation are controlled by covalent bonds and nonbonded interactions, operating through space between neighboring groups of atoms or local charges within the molecules.[1]

Weinkauf **et al.** suggested that charge localization in the peptide be achieved by resonant UV two-photon ionization at an aromatic chromophore.[2] In other words, resonant absorption of a first visible photon activates charge flow in the cations and that of a second one probes presence of charge in the aromatic chromophore.

According to ref 3, a surplus charge in flexible molecules can influence the conformational landscape. For the reason, there exists correlation among structure, charge and reactivity of biomolecules, namely amino acids and proteins.

A new grouping of the stable conformers into two subgroups I and II is given based on the calculated results together with the results of the charge distribution in cations described in section 6.3. Delocalization of charge distributions between the amino and phenyl groups in L-phenylalanine on ionization is explained in terms of "hyperconjugation" in which the highest occupied molecular orbital (HOMO) of the phenyl group and the nonbonding molecular orbital (n-MO) of the

amino group form new bonding and antibonding MOs.

The charge distribution for cations is obtained using natural bond orbital[4] (NBO) analysis. The charge distribution between the subgroups of L-tyrosine and L-tryptophan are investigated in addition to L-phenylalanine.

Accurate evaluation of conformation-dependent charge distributions is very important in studying the mechanisms of ionization-induced chemical reactions as well as determination of conformers.[3,5-7]

This chapter explains the different trend of charge distribution among the three aromatic amino acids by using the ab initio MO method[8] and DFT method at the M05-2X level of theory[9-11] taking into account the effect of noncovalent, long-range potential.

6.2. Computational procedure

All the calculations are carried out including zero-point energy corrections (ZPE_{corr}). Each conformer for all the three amino acids is optimized in the corresponding neutral ground state by setting the charge and the spin multiplicity to 0 and 1, respectively at M05-2X/6-311+G* level of theory. Their optimized geometries in neutral ground state are used as initial geometry for each cationic conformer based on Franck-Condon principle and then they are reoptimized by setting the charge and the spin multiplicity to 1 and 2, respectively, at M05-2X with 6-311+G* basis set in the same way as geometry optimization for the neutral conformers.

For investigation of their charge distribution, natural bond orbital (NBO) analysis is carried out. The charge analysis can be obtained by subtraction between the partial charges of cation and neutral, and the value of charge distribution from M05-2X result can be shown using GaussView5 without difficulty. The checkpoint file (*.chk), which is created as normal termination, should be converted to its formatted checkpoint file (*.fchk) in order to visualize the shape of molecular orbital for the optimized structure in the ground state.[4] B3LYP calculation also were accomplished with 6-311+G* basis set for comparison with the M05-2X results.

6.3. Results and discussion

In Figure 6.1, toluene, cresol, and skatol are employed as representative aromatic residue chromophores for L-phenylalanine, L-tyrosine, and L-tryptophan, respectively.

A typical molecule for charge doubly localization (division) upon ionization is 2-phenylethylamine[3] (2-PEA), analogue of phenylalanine, which is supposed to be formed by ethylbenzene as a residue and ethylamine as a backbone. (See Figure 6.1) According to ref 3, they have almost the same ionization energy (8.77 eV for ethylbenzene, 8.80 eV for ethylamine) in experiment .

Weinkauf **et al**. suggested that the charge be founded to be delocalized equally over the two sites in the case of similar ionization energies and the charge delocalization effect can be observed through “bond” or “space”.

For the chromophores of those aromatic amino acids, we reported that only toluene and alanine belonging to subgroup II have the similar adiabatic ionization energy.[6,7] In brief, the adiabatic ionization energies by M05-2X are 9.36, (8.94), 8.74, 8.12, and 7.53 eV for alanine belonging to subgroup I (II), toluene, cresol, and skatol, respectively.

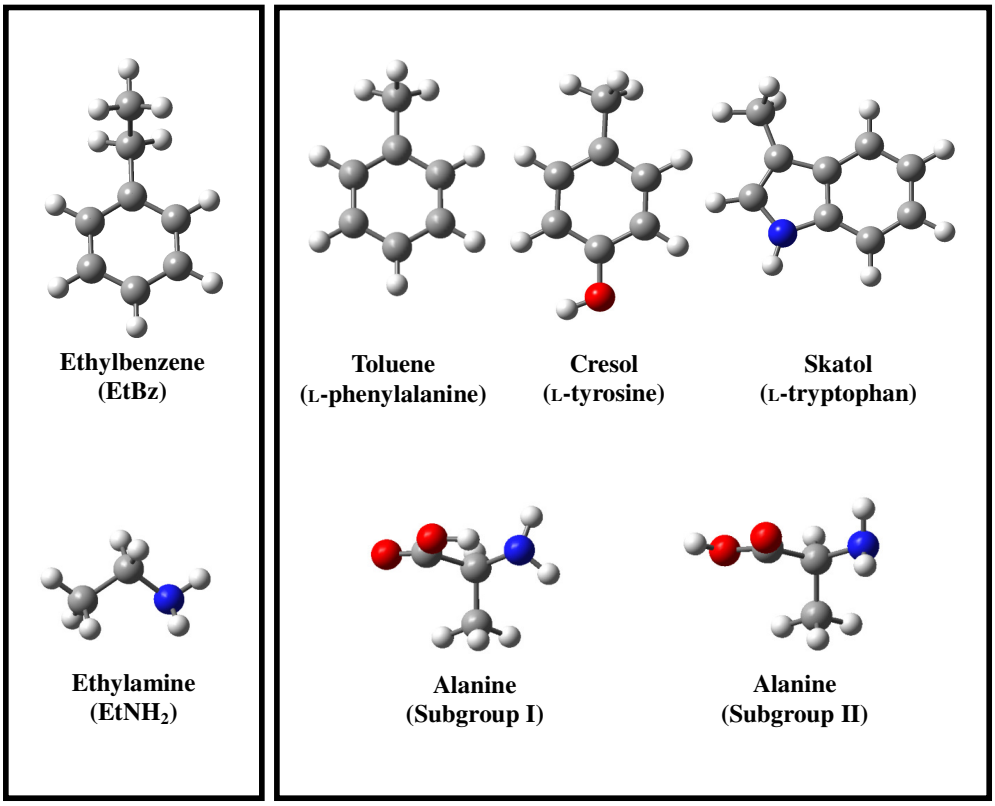


Figure 6.1. Several chromophores for explanation of charge distribution of the three aromatic amino acids.

6.3.1. L-phenylalanine

As mentioned above, remind the correlation between structure and charge of amino acids. Conformers of L-phenylalanine belonging to a given subgroup share the same characteristic features in the cationic ground state, that is, their structure for subgroup II have amino group's planarity, the elongated C_α–C_β bond length, and the cationic charge delocalized at two sites.

Whereas, those of subgroup I have the pyramidal amino group, normal C_α–C_β bond length, and the cationic charge localized in one region. A theoretical support for these findings has been recently given on the basis of M05-2X calculations.[6]

Table 6.1 shows the M05-2X results for the charge distributions on two chromophores, the phenyl ring and amino groups of the nine L-phenylalanine conformers. For comparison, B3LYP results also are listed in parentheses. Both the M05-2X and B3LYP results indicate that the charge distributions of conformers belonging to subgroup I are mostly localized on the phenyl ring.

However, for subgroup II conformers except conformer V, they are qualitatively delocalized, that is, the charges are shared between the two chromophores. The newly calculated value of 0.81 for conformer V strongly suggests that the cationic conformer belongs to subgroup I. In addition, the newly obtained results for the charge distributions show remarkable contrast in charge

distributions between the two subgroups compared with those obtained by the B3LYP method. In chapter 4, conformer V on ionization has already been verified as a hybrid type conformer based on structural change from M05-2X result.

In particular, charges of subgroup II conformers are almost equally distributed between the phenyl and the amine groups. The differences in charge distributions between the two subgroups were explained by the form of their intramolecular hydrogen bonding:[12] $-COOH \rightarrow -NH_2$ vs. $-NH_2 \rightarrow -OCOH$.

In addition to intramolecular hydrogen bonding ($-COOH \rightarrow -NH_2$) for conformers [I(X) and III(B)] belonging to subgroup I, we should stress that hyperconjugation in conformers [VII(A), VI(C), II(D), and IX(E)] belonging to subgroup II is responsible for the delocalized charges as qualitatively indicated in Table 6.1.

Figure 6.2 shows the phases of HOMO and HOMO-1 of the observed conformers: It can be seen that for subgroup II hyperconjugative HOMOs are formed through space interactions between the n-MO of the amino group and the π orbital of the phenyl ring [VII(A), IX(E)] or through bond interactions [(VI(C), II(D)].

Table 6.1. Partial charge distributions of cationic L-phenylalanine conformers.

Subgroup	Experimental Assignment	Conformer	M05-2X (B3LYP)			
			Phenyl group		Amino group	
I	X	I	0.79	(0.62)	0.04	(0.09)
I	B	III	0.73	(0.53)	0.09	(0.17)
I		IV	0.73	(0.54)	0.09	(0.16)
II (I)		**V**	**0.81**	**(0.42)**	**0.06**	**(0.33)**
II	A	VII	0.34	(0.41)	0.34	(0.35)
II	C	VI	0.38	(0.44)	0.33	(0.34)
II	D	II	0.39	(0.44)	0.31	(0.34)
II	E	IX	0.34	(0.41)	0.34	(0.35)
II		VIII	0.33	(0.40)	0.34	(0.35)

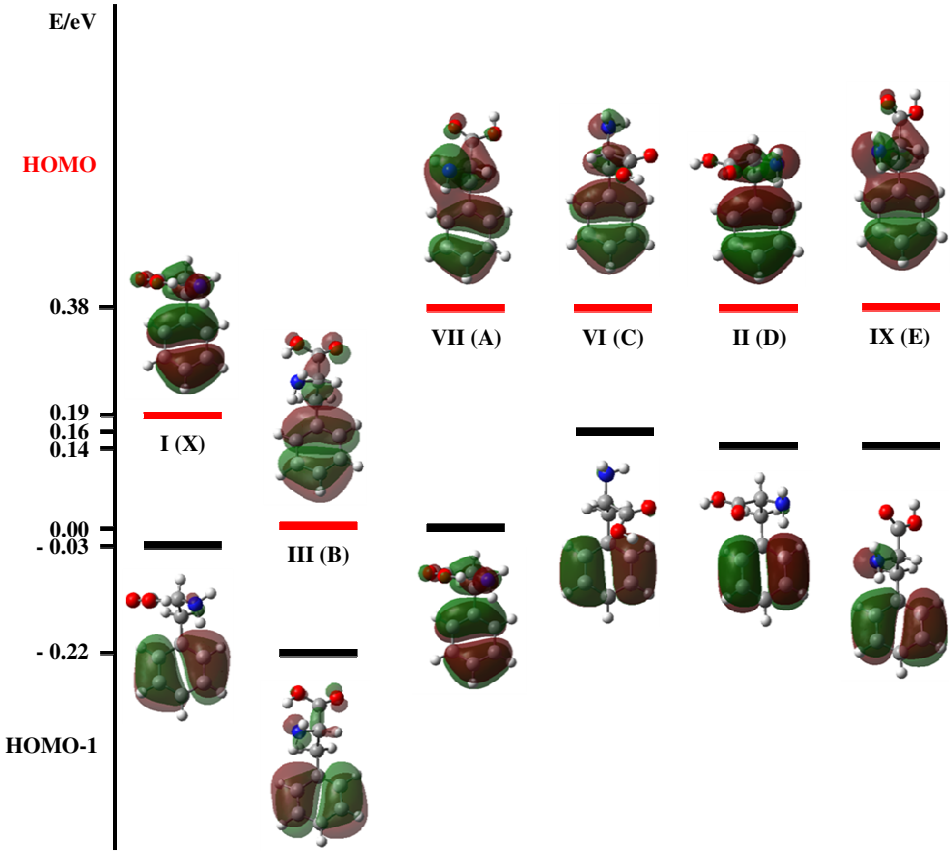


Figure 6.2. Molecular orbital phases and energies for HOMOs and HOMO-1s of six L-phenylalanine conformers.

6.3.2. L-tyrosine and L-tryptophan

Results of charge distribution obtained by natural bond orbital (NBO) analysis are given in Table 6.2. a(b) for L-tyrosine and L-tryptophan, respectively. It can be seen from the M05-2X results that their charge of 90% is localized on the residue (phenol and indole) for all of the L-tyrosine and L-tryptophan conformers belonging to subgroup I. The B3LYP results are marked in parentheses.

Note their charge localization for all the conformers belonging to subgroup II of L-tyrosine and L-tryptophan.[7] The values of their partial charge distribution are definitely equal to those of conformers belonging to subgroup I.

This makes clear contrast with L-phenylalanine conformers belonging to subgroup II, in which charge is doubly localized (divided) in each phenyl group and amino group with equal weight.[6] It is considered that the different tendency for charge localization between L-phenylalanine vs. L-tyrosine (L-tryptophan) is concerned with their adiabatic ionization energies predicted in the previous chapter.

Table 6.2. Partial charge distributions of L-tyrosine and L-tryptophan conformers in cations.

(a) L-tyrosine

Subgroup	Experiment Assignments	Conformer	M05-2X (B3LYP)			
			Phenol group		**Amino group**	
I	14	I_l	0.89	(0.78)	0.03	(0.07)
I	15	I_r	0.89	(0.78)	0.03	(0.08)
I	6	III_l	0.85	(0.69)	0.04	(0.10)
I	4	III_r	0.85	(0.69)	0.04	(0.10)
II	16	II_l	0.90	(0.75)	0.02	(0.10)
II	17	II_r	0.90	(0.75)	0.02	(0.09)
II	7	VII_l	0.91	(0.67)	0.01	(0.17)
II	5	VII_r	0.91	(0.65)	0.01	(0.18)

(b) L-tryptophan

Subgroup	Experiment Assignments	Conformers	M05-2X (B3LYP)			
			Indole group		**Amino group**	
I	A	I_a	0.90	(0.85)	0.02	(0.04)
I	F	III_a	0.87	(0.81)	0.04	(0.05)
II	E	II_a	0.89	(0.84)	0.03	(0.06)
II	C	II_b	0.92	(0.90)	0.01	(0.01)
II	B	V_a	0.90	(0.84)	0.03	(0.06)
II	D	V_b	0.93	(0.92)	0.01	(0.01)

Let us now make a semiquantitative discussion about the charge distribution shown in Table 6.2a(b) by using a simple two-state model introduced in Figure 5.4. According to the model, it consists of the HOMO of aromatic residue, ϕ_π, and nitrogen nonbonding MO of the backbone, ϕ_n. The HOMO (MO from which electron is released) of the aromatic amino acid $\Phi_n$can be expressed by a linear combination of these two MO's as $\Phi = \beta\,\phi_\pi - \alpha\phi_{nN}$. α and β are coefficients given as

$$\alpha = \frac{\left\{(\Delta\varepsilon^2 + 4\gamma^2)^{\frac{1}{2}} - \Delta\varepsilon\right\}}{\left[\left\{(\Delta\varepsilon^2 + 4\gamma^2)^{\frac{1}{2}} - \Delta\varepsilon\right\}^2 + 4\gamma^2\right]^{\frac{1}{2}}}$$

and

$$\beta = \frac{2|\gamma|}{\left[\left\{(\Delta\varepsilon^2 + 4\gamma^2)^{\frac{1}{2}} - \Delta\varepsilon\right\}^2 + 4\gamma^2\right]^{\frac{1}{2}}} \quad .$$

Here, $\Delta\varepsilon = \varepsilon_\pi - \varepsilon_{nN}$ and $\gamma = \langle\, \phi_\pi \big| \hat{H} \big| \phi_{nN} \rangle$; γ is the interaction energy between ϕ_π and ϕ_{nN}. $\hat{H}$ is the electronic Hamiltonian of aromatic amino acid at the ground state geometrical structure.

For L-phenylalanine belonging to subgroup II, $|\alpha| = |\beta|$ since both $\phi_\pi$and ϕ_{nN} have the same value of orbital energy and $\Delta\varepsilon = 0$. This means that charge is delocalized (divided) between the backbone and aromatic residue with equal weight, i.e., $|\alpha|^2 = |\beta|^2$.

In this case, γ can be estimated as the ionization energy difference between phenyl (or amino) group and L-phenylalanine. γ is estimated to be 0.32 eV from the experimental ionization energies.

In sequence, let us consider charge distribution of L-tyrosine by using the two-state model in Figure 5.4. Coefficients α and β were estimated to be $|\alpha|^2 = 0.11$ and $|\beta|^2 = 0.89$. Here, we adopt $\gamma = 0.32$ eV, which is estimated from L-phenylalanine since the aromatic ring of L-tyrosine has the similar aromatic ring as that of L-phenylalanine. It should be noted from comparison of geometrical structures between L-tyrosine and L-phenylalanine that their conformers have one-to-one correspondence with respect to their conformation.[7]

Furthermore, the angle between the directions of two orbitals ($\phi_\pi$and ϕ_{nN}) of each L-tyrosine conformer is almost equal to that of the corresponding L-phenylalanine conformer. Here, $\Delta\varepsilon$ is calculated as 0.82 eV and it is taken from the theoretical ionization energies of two substructures, cresol and alanine belonging to subgroup II. The values of the charge distribution obtained by using the simple model semiquantitatively explain those shown in Table 6.2(a).

In the similar way, we obtain $|\alpha|^2 = 0.44$ and $|\beta|^2 = 0.96$ for L-tryptophan with $\Delta\varepsilon = 1.41$ eV. It can be seen that there is significant difference in the charge distribution between the *ab* initio calculation and the simple two-state model. The main difference originates from use of the same interaction energy $\gamma = 0.32$ eV as that used in L-phenylalanine for evaluation of L-tryptophan charge distribution

in the simple model even though there is a large discrepancy in aromatic rings between L-tryptophan and L-phenylalanine. Actually, $\gamma = 0.53\text{eV}$ is obtained from the expression for $\alpha(\beta)$ with *ab* initio values of the charge distribution for L-tryptophan in Table 6.2(b).

On the other hand, $\gamma = 0.31$ eV is obtained from the expression for $\alpha(\beta)$ with *ab* initio values of the charge distribution for L-tyrosine in Table 6.2(a) and there is almost no discrepancy of γ between L-phenylalanine and L-tyrosine.

In summary, charge delocalization for the subgroup II conformers of L-phenylalanine can be considered as a special case. The tendency of no C_α–C_β elongation and nonplanarity of the C_α–NH_2 for all of the L-tyrosine and L-tryptophan conformers in radical cations, as shown in Tables 4.3 and 4.5, can be qualitatively explained using the same model.

Charge localization of aromatic residue in aromatic amino acids means that ionization takes place from the HOMO of aromatic residue. Therefore, the degree of contribution of molecular orbital character like nN and π plays an important role in determining whether the charge is localized or delocalized.

6.4. Conclusion

Conformation-dependent properties of the aromatic amino acids L-tyrosine and L-tryptophan against L-phenylalanine in radical cations have been studied by using new density functional theory, M05-2X. The results showed that all of the conformers of both L-tyrosine and L-tryptophan undergo charge localization respective of formation of intramolecular hydrogen bonding.

Such localization behavior originates from the existence of HOMO energy difference between the backbone and residue of the aromatic amino acid. Among the three aromatic amino acids, L-phenylalanine, L-tyrosine, and L-tryptophan, only subgroup II conformers of L-phenylalanine without intramolecular hydrogen bonding (–COOH → $–NH_2$) in the backbone undergo charge delocalization.

This is due to nearly the same energy values of HOMO's between the backbone, alanine, and aromatic residue, toluene. The analysis given here provides clues to analyze the dynamics of charges in peptides and amino acids.

6.5. References

[1] Robertson, E. G.' and Simons, J. P. *Phys. Chem. Chem. Phys.*, **2001**, *3*, 1

[2] Weinkauf, R.; Schanen, P.; Metsala, A.; Schlag, E. W.; Burgle, M.; Kessler, H. *J. Phys. Chem.* **1996**, *100*, 18567.

[3] Weinkauf, R.; Lehrer, F.; Schlag, E. W.; Metsala, A. *Faraday Discuss.* **2000**, *115*, 363.

[4] Gaussian 09, Revision A.1, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc.,

Wallingford CT, **2009**.

[5] Lee, K. T.; Kim, H. M.; Han, K. Y.; Sung, J.; Lee, K. J.; Kim, S. K. *J. Am. Chem. Soc.* **2007**, *129*, 2588.

[6] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583.

[7] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2011**, *115*, 9658.

[8] Sahnoun, R.; Fujimura, Y.; Kabuto, K.; Takeuchi, Y.; Noyori, R. *J. Org. Chem.* **2007**, *72,* 7923.

[9] Zhao, Y; Truhlar, D. G. *J. Phys. Chem. A* **2004,***108,* 6908.

[10] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[11] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[12] Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321,* 49.

국문 초록

원거리 비공유결합 상호작용인 분자 내 수소결합을 하는 분자 시스템에 최적화된 새로운 밀도함수이론 (density functional theory) 방법인 M05-2X 을 이용하여 중성 및 양이온 상태에서의 세 가지 방향족 아미노산 (페닐알라닌, 타이로신, 트립토판)의 구조-의존적 특성을 연구하였다.

우선, M05-2X 방법의 신뢰도를 입증하기 위해서 몇 가지 발색단에 대하여 이 방법을 적용하였다. 또한 M05-2X 계산 결과와 비교하기 위해서 보편적으로 사용되는 밀도함수이론 방법인 B3LYP 계산도 병행하였다. M05-2X 로부터 예측된 이론값과 실험적으로 관측된 값은 오차범위 내에서 일치하였으나 B3LYP 의 계산결과는 실험치와 상당한 차이가 있음을 보였다. M05-2X 은 이들 발색단 중에서도 특히 방향족 아미노산에 대한 이온화 에너지 계산에서 실험치와 정확히 일치함을 보였는데, 이 결과는 M05-2X 가 분자 내 수소결합을 하는 생체 분자의 이론적 연구에 있어 가장 적합한 방법이라는 것을 증명한 것이라 하겠다.

본 연구에서의 모든 계산은 두 가지 계산방법 (M05-2X/6-311+G*와 B3LYP/6-311+G*)을 모두 적용하였으며, 영점 에너지 보정을 하였다. 중성과 양이온의 바닥상태에서 세 가지 방향족 아미노산의 구조 이성질체들에 대한 구조 최적화와 이온화 에너지, 전하 분포 계산을 수행하였다. 이론적으로는 예측 가능하지만 실험적으로 관측되지 않는 페닐알라닌의 몇몇 구조 이성질체에 대해 구조 이성질화 과정을 고유 반응 좌표 (intrinsic reaction coordinate) 경로로부터 규명하였다.

세 가지 방향족 아미노산의 구조 이성질체에 대해 M05-2X 는 그들의 실험적인 단열 이온화 에너지(adiabatic ionization energy)와 수직 이온화 에너지(vertical ionization energy)를 성공적으로 재현하는 반면, B3LYP 은 실험적으로 관찰된 값에 비해 0.2 ~ 0.3 eV 만큼 상당히 낮은 에너지를

산출하였다. 양이온 상태에서 그들의 구조 이성질체들에 대한 전하 분포를 자연적 결합 궤도 (natural bond orbital) 분석을 이용하여 조사하였다. 페닐알라닌의 구조 이성질체는 분자 내 수소결합의 형태에 따라 전하가 편재화 되거나 비편재화 되지만, 타이로신과 트립토판의 모든 구조 이성질체들은 그들이 이온화 되었을 때 분자 내 수소 결합의 형태와 무관하게 전하 편재화만 일어난다는 사실을 밝혀냈다.

세 가지 방향족 아미노산 가운데 오직 페닐알라닌에 대해서만 전하 비편재화가 일어나는 원리를 규명하기 위하여 간단한 이론 모델을 제시하였다. 그 모델의 기본 개념은 이들 방향족 아미노산이 내부적으로 두 개의 별도의 양이온 코어를 갖는 골격과 잔기로 구성된다는 가정에서 출발한다. 두 개의 양이온 코어 사이의 단열적 이온화 에너지의 차이가 전하 편재화의 정도를 결정한다는 사실을 확인하였다.

세 가지 방향족 아미노산은 생체 분자 내에서 발색단 역할을 하기 때문에 이들 분자의 전하 분포에 대한 분석 결과는 단백질이나 다른 아미노산에서의 전하 동력학을 해석하는데 있어서 중요한 단서를 제공할 것이다.

본 연구에서 다뤄진 세 가지 방향족 아미노산의 구조-의존적 특성과 이온화 에너지 사이에 밀접한 관련이 있음을 이론적으로 규명하였다.

주요어: M05-2X, 비공유결합 원거리 효과, 페닐알라닌, 타이로신과 트립토판, 전하 편재화, 이온화 에너지

학번: 2002-30939

理學博士學位論文

Theoretical Study on Conformation-Dependent Properties of the Three Aromatic Amino Acids

세 가지 방향족 아미노산의 구조-의존적 특성에 관한 이론적 연구

2014 年 2 月

서울大學校 大學院
化學部 物理化學專攻
白 炅 燁

Ph.D. Dissertation

Theoretical Study on Conformation-Dependent Properties of the Three Aromatic Amino Acids

February 2014

Supervisor: Professor Seong Keun Kim

Co-supervisors: Professor Dr. Sheng Hsien Lin

Professor Dr. Yuichi Fujimura

Department of Chemistry
Seoul National University
Kyung Yup Baek

Abstract

Theoretical Study on Conformation-Dependent Properties of the Three Aromatic Amino Acids

Kyung Yup Baek

Department of Chemistry

The Graduate School

Seoul National University

Study on conformation-dependent properties of the three aromatic amino acids (L-phenylalanine, L-tyrosine and L-tryptophan) in neutral and radical cations has been performed by using density functional theory (DFT) with a new density functional M05-2X, which is applicable to molecular systems with noncovalent long-range interactions such as intramolecular hydrogen bonding.

First of all, M05-2X is applied to several simple chromophores in order to verify reliability of the method as well. The results obtained from M05-2X are compared with those estimated by using the conventional DFT (B3LYP) and

can explain the correspondence between the observed and predicted chromophores without ambiguity. On the other hand, a considerable discrepancy between the predicted results from DFT (B3LYP) and experimental ones is found. This suggests that noncovalent long-range effects should be included for accuracy of calculation.

Optimized geometrical structures for both the neutral and cationic conformers of the amino acids are evaluated with those DFT methods and the procedure of conformerization for L-phenylalanine conformers in neutral is indicated from the results of intrinsic reaction coordinate (IRC) profile.

As described above, M05-2X method successfully reproduces experimental adiabatic and vertical ionization energies for all the three aromatic amino acids, whereas B3LYP functional yields significantly lower ionization energies from the observed ones by 0.2 ~ 0.3 eV.

Charge distributions in the cationic conformers are estimated by using natural bond orbital (NBO) analysis. All the conformers of L-tyrosine and L-tryptophan become charge localization when they are ionized regardless of the type of intramolecular hydrogen bonding, unlike the case of L-phenylalanine that was clarified earlier by other studies.

Their charge distribution analysis helps dynamics of charges in proteins to be understood because the three aromatic amino acids play a role as chromophores in biomolecules.

A simple model is employed in order to figure out different degrees of charge localization among all the three aromatic amino acids. The basic concept of the model starts from postulation that the aromatic amino acid consists of two sub-moieties of distinct cationic core: the backbone and aromatic residue. The difference in adiabatic ionization energy between these two sub-moieties is found to govern the degree of charge localization.

Here it is illuminated theoretically that there exists considerable correlation between conformation-dependent property and ionization energy for the three aromatic amino acids.

Keywords: M05-2X, noncovalent long-range effect, L-phenylalanine, L-tyrosine and L-tryptophan, charge localization, ionization energy

Student Number: 2002-30939

Contents

Chapter 1.

Introduction

In 1900, Max Planck offered a radical proposal that blackbody radiation emitted by microscopic particles was limited to certain discrete values, i.e., it was 'quantized'.[1] Such quantization was essential to reconciling large differences between predictions from classical models and experimental results.

As the twentieth century progressed, it became increasingly clear that quantization was not only a characteristic of light, but also of the fundamental particles from which matter is constructed. Bound electrons in atoms, in particular, are clearly limited to discrete energies (levels) as indicated by their ultraviolet and visible line spectra.

This phenomenon has no classical correspondence – in a classical system, obeying Newtonian mechanics, energy can vary continuously. In order to describe microscopic systems, another different mechanics was required.

Matter at the atomic and molecular or microscopic level reveals the existence of a variety of particles which are identifiable by their distinct properties, such as mass, charge, spin, and magnetic moment. All of these seem to be of a quantum nature in the sense that they take on only certain discrete values. This discreteness of physical properties persists when particles combine to form nuclei, atoms, and molecules.[2]

The development of quantum mechanical techniques are more generally applicable and can be implemented on a computer. Now, quantum mechanics can be used to perform calculations molecular systems of real and practical interest.[3]

Computational chemistry which attained general currency as the same meaning of theoretical chemistry is based on molecular quantum mechanics with Schrödinger equation. It has been great tool which is employed to understand molecular properties to the motion and interactions of electrons and nuclei.[4]

Actually, any chemical species (unstable conformers and other reactive intermediates in neutral) which might be difficult to investigate experimentally may be predicted theoretically.[5]

Over the last century, there have been accelerated advances in the field of computational chemistry and they have enabled development of efficient quantum chemical calculation methods employing electronic structure theories such as DFT (B3LYP)[6-8] and a new DFT functional M05-2X.[9,10]

These theoretical methods are essential for many studies on optimized geometry and electronic properties. Nowadays, combinations of supercomputers with high-performance and Gaussian software package which is one of enhanced electronic structure programs are extremely useful to interpret complicated experimental data.

1.1. The three aromatic amino acids

Because of importance as building blocks of proteins and roles in biomolecules, many studies on amino acids have been done through laser spectroscopic techniques combined with quantum chemical calculation methods.[5,11-19]

The three aromatic amino acids (Figure 1.1) can be divided into two substructure, backbone of α-amino acid like alanine and aromatic residue like toluene, cresol and skatol. L-phenylalanine and L-tyrosine have same structures except –OH group attached to aromatic ring.

It has been shown that the aromatic amino acids have many conformers depending on internal charge distributions and external conditions.[5,11-14] It was reported that L-phenylalanine is biologically converted into L-tyrosine in the liver by phenylalanine hydroxylase and L-tyrosine is a precursor to neurotransmitter dopamine which is known to be related to Parkinson's disease.[20,21]

On the other hands, L-tryptophan is known to be a precursor for serotonin which is a neurotransmitter modulating central and peripheral functions through action on neurons, smooth muscle, and other cell types.[22,23]

It was reported that tryptophan depletion consequently brings to significant reduction serotonin (5-HT) synthesis and release in brain and lack of serotonin recurs symptoms in patients with depression and panic disorder, who have responded to treatment with antidepressant due to reduction.[24]

Since chromophores of these three aromatic amino acids have high optical absorption cross sections at 250 – 285 nm in the ultraviolet (UV) region,[25] their spectroscopic studies have been used as optical probes for structures and dynamics of proteins.[26]

Aromatic amino acids have also been utilized as a trigger to charge transfer dynamics of polypeptides.[14,27,28]

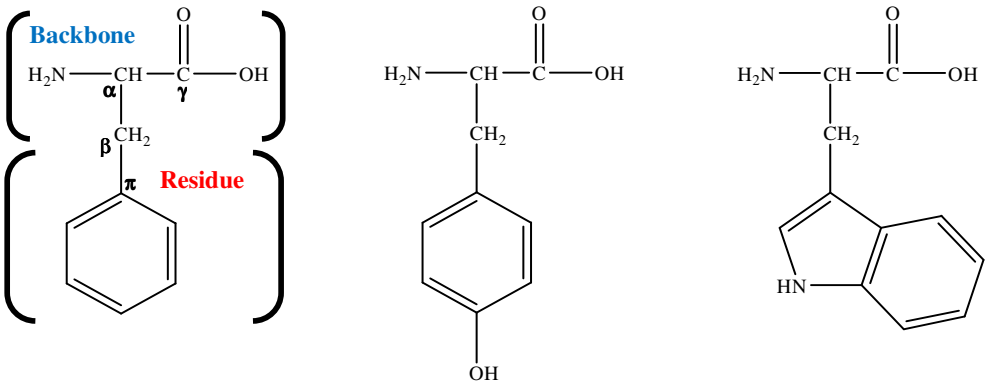


Figure 1.1. Molecular structure of L-phenylalanine, L-tyrosine and L-tryptophan.

1.2. Intramolecular interactions

A hydrogen bond is a kind of typical noncovalent bond, that is, it is not direct covalent bond but electromagnetic attractive interaction between polar atoms (molecules). This interaction is called “intermolecular” hydrogen bonding when it is occurred between molecules, whereas it is called “intramolecular” hydrogen bonding when it is occurred within different parts of a single molecule.

According to types of intramolecular hydrogen bonding, all conformers of the aromatic amino acids were known to be classified into two subgroups, I and II.[29] (Subgroup I characterized by $-COOH \rightarrow -NH_2$ vs. subgroup II by $-NH_2 \rightarrow -OCOH$ in Figure 1.2.)

Noncovalent long-range interactions between a partially charged backbone and a residue in the amino acids play a crucial role in determining stable conformers, especially, ionic conformers.[14] B3LYP has been considered one of the most famous functional. However, it involves some critical defects due to absence of noncovalent interactions in the functional.[30]

Recently, new functionals including the noncovalent effect have been developed in order to complement the shortcoming of B3LYP, which show good agreement with experiments than B3LYP.[30-36]

For some chromophores, we carried out quantum chemical calculation of geometry optimization and estimated their adiabatic ionization energies based on M05-2X in order to investigate how much long-range effects impact on the accuracy of the results.

Especially, the M05-2X results of 2-phenylethylamine (2-PEA) and L-phenylalanine with intramoelcular hydrogen bonding show good agreement with the experimental data compared with the B3LYP results as shown in Figure 1.3.

This indicates that it is essential to apply a new DFT functional M05-2X to neutrals and, in particular, cationic conformers of amino acids in order to ascertain correspondence between observed and predicted conformers.

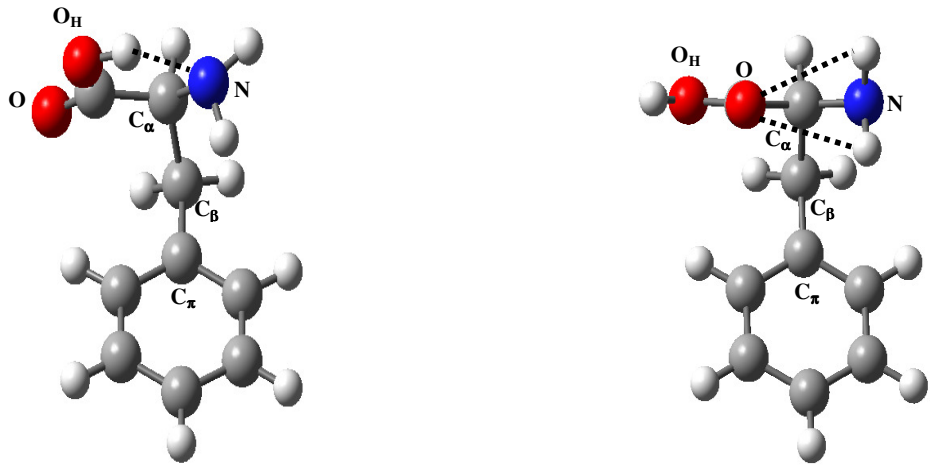


Figure 1.2. Representative structure of subgroup I and II in aromatic amino acid

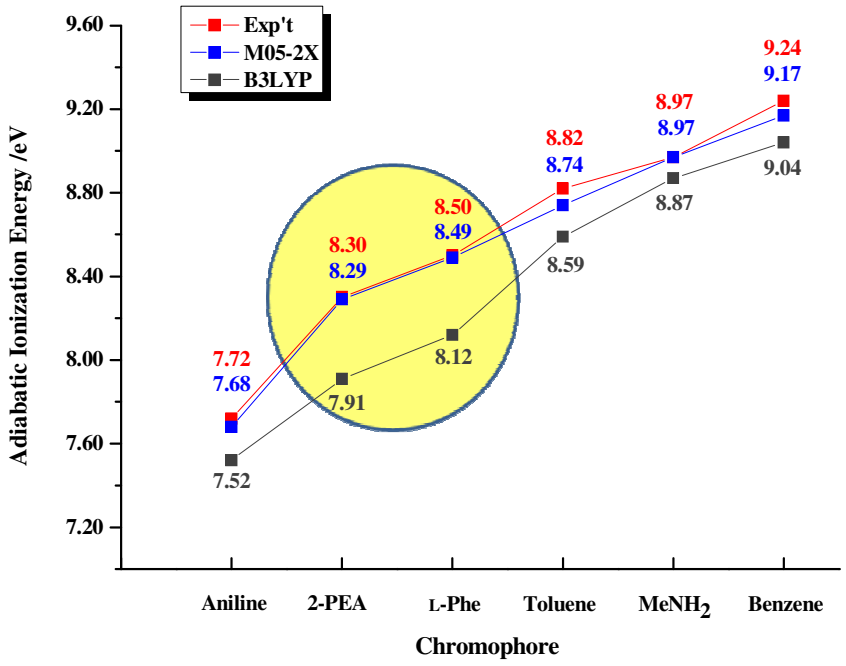


Figure 1.3. Adiabatic ionization energies for several chromophores.

1.3. Charge localization vs. delocalization

Let us consider ionized species of a molecule that consists of two moieties such as residue (A) and backbone (B), respectively as shown in Figure 1.4. There exist two cases depending on whether moieties have same ionization energy or not.

It has been reported that degree of charge localization is proportional to ionization energy gap between two moieties of molecules and 2-PEA, which consists of two moieties (ethylamine and ethylbenzene), were employed as a reference molecule to explain charge delocalization due to be equal to ionization energies each other.[11,37]

This approach is useful to interpret the different tendencies of charge distribution (localization or delocalization) for the three aromatic amino acids (L-phenylalanine, L-tyrosine, and L-tryptophan), and to investigate conformation-dependent properties between the subgroups (subgroup I and II) even though they belong to the same molecule.[5,11]

In the thesis, main topic is M05-2X results on conformation-dependent properties such as geometrical structure, adiabatic (vertical) ionization energies, charge distributions for the neutral (cationic) conformers of these aromatic amino acids. In addition, a simple molecular orbital (MO) model is introduced to explain the different tendency of their ionization energies and charge distributions.

In the next two chapters, a theoretical background relevant to all these work and computational method we have employed is presented.

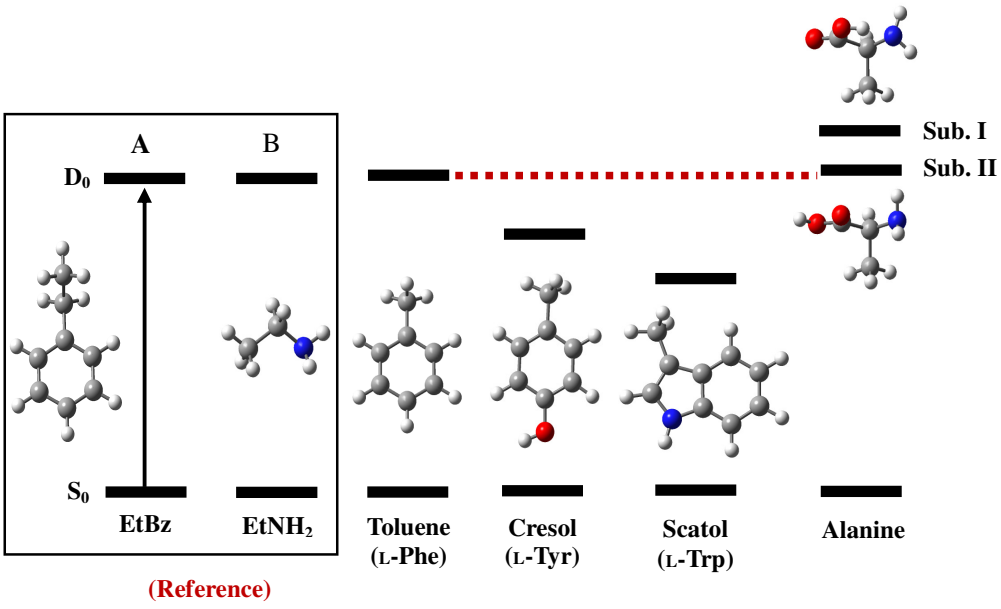


Figure 1.4. Ionized backbones and residues corresponding aromatic amino acids.

1.4. References

[1] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), p. 95.

[2] Eugen Merzbacher, *Quantum Mechanics* (John Wiley & Sons, Ltd., New York, 1998), p. 1.

[3] Andrew R. Leach, *Molecular Modelling* (Anddison Wesley Longman Ltd, Harlow Essex, 1996), p. 25.

[4] Warren J. Hehre; Leo Radom.; Paul v. R. Schleer; John A. Pople, *Ab initio Molecular Orbital Theory* (John wiley & Sons Inc., New York, 1986), p. 1.

[5] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583.

[6] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

[7] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[8] Lee, C.; Yang, W.; Parr, R.G. *Phys. Rev. B* **1988**, *37*, 785.

[9] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[10] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[11] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2011**, *115*, 9658.

[12] Lee, K. T.; Kim, H. M.; Han, K. Y.; Sung, J.; Lee, K. J.; Kim, S. K. *J. Am. Chem. Soc.* **2007**, *129*, 2588.

[13] Lee, K. T.; Sung, J.; Lee, K. J.; Kim, S. K.; Park, Y. D. *Chem. Phys. Lett.* **2003**, *368,* 262.

[14] Weinkauf, R.; Schanen, P.; Yang, D.; Soukara, S.; Schlag, E. W. *J. Phys. Chem.* **1995**, *99*, 11255.

[15] Robertson, E. G.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3,* 1.

[16] Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321,* 49.

[17] Hashimoto, T.; Takasu, Y.; Yamada, Y.; Ebata, T. *Chem. Phys. Lett.* **2006**, *421,* 227.

[18] Lee, Y.; Jung, J.; Kim, B.; Butz, P.; Snoek, L. C.; Kroemer, R. T.; Simons, J. P. *J. Phys. Chem. A* **2004**, *108,* 69.

[19] Huang, Z.; Yu, W.; Lin, Z., *Theochem-J. Mol. Struct.* **2006**, *758,* 195.

[20] Elsworth, J. D.; Roth, H. R. *Exp. Neurol.* **1997**, *144*, 4.

[21] Grace, L. I.; Cohen, R.; Dunn, T. M.; Lubman, D. M.; de Vries, M. S. *J. Mol. Spectrosc.* **2002**, 2*15*, 204.

[22] Shi, Z.; Devasagayaraj, A.; Gu, K.; Jin, H.; Marinelli, B.; Samala, L.; Scott, S.; Stouch, T.; Tunoori, A.; Wang, Y.; Zang, Y.; Zhang, C.; Kimball, S. D.; Main, A. J.; Yu, X.; Buxton, E.; Patel, S.; Nguyen, N.; Swaffield, J.; Powell, D. R.; Wilson, A.; Liu, Q. *J. Med. Chem.* **2008**, *51*, 3684.

[23] Lima, S.; Kumar, S.; Gawandi, V.; Momany, C.; Phillips, R. S. *J. Med. Chem.* **2009**, *52*, 389.

[24] Bell, C.; Abrams, J.; Nutt, D. *Brit. J. Psychiat.* **2001**, *178*, 399.

[25] Steiner, R. F.; Weinryb, I. *Excited States of Proteins and Nucleic Acids* (Macmillan, London, 1971).

[26] Beechem, J. M.; Brand, L. *Annu. Rev. Biochem.* **1985**, *54*, 43.

[27] Weinkauf, R.; Schanen, P.; Metsala, A.; Schlag, E. W.; Burgle, M.; Kessler, H. *J. Phys. Chem.* **1996**, *100*, 18567.

[28] Lehr, L.; Horneff, T.; Weinkauf, R.; Schlag, E. W. *J. Phys. Chem. A* **2005**, *109*, 8074.

[29] Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321,* 49.

[30] Zhao, Y.; Truhlar, D. G. *Accounts Chem. Res.* **2008**, *41,* 157.

[31] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[32] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[33] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123,* 161103.

[34] Suresh, C. H.; Mohan, N.; Vijayalakshmi, K. P.; George, R.; Mathew, J. M. *J. Comput. Chem.* **2008**, *30,* 1392.

[35] Sato, T.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2007,** *126,* 234114.

[36] Hohenstein, E. G.; Chill, S. T.; Sherrill, C. D. *J. Chem. Theory Comput.* **2008,** *4,* 1996.

[37] Weinkauf, R.; Lehrer, F.; Schlag, E. W.; Metsala, A. *Faraday Discuss.* **2000**, *115*, 363.

Chapter 2.

Background

2.1. Schrödinger equation

The fundamental postulation of quantum mechanics is that a so-called wave function, Ψ, exists for any (chemical) system, and that appropriate operators which act upon Ψ generate the observable properties of the system.

Thus, we write

$$H\Psi = E\Psi\,, \tag{2.1}$$

which is the Schrödinger equation.[1] The operator in Eq. (2.1) that returns the system energy, E as an eigenvalue is called the Hamiltonian operator, H. The typical form of the Hamiltonian operator with which we will be concerned takes into account five contributions to the total energy of a system: the kinetic energies of the electrons and nuclei, the attraction of the electrons to the nuclei, and the interelectronic and internuclear repulsions.

In more complicated situations, e.g., in the presence of an external electric field, in the presence of an external magnetic field, in the event of significant spin-orbit coupling in heavy elements, taking account of relativistic effects, etc., other terms are required in the Hamiltonian. Casting the Hamiltonian into mathematical notation, we have

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i \sum_k \frac{e^2 \mathrm{Z}_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 \mathrm{Z}_k \mathrm{Z}_l}{r_{kl}}\,, \tag{2.2}$$

where i and j run over electrons, k and l run over nuclei, $\hbar$ is Planck's

constant divided by 2π, m_e is the mass of the electron, m_k is the mass of nucleus k, ∇^2 is the Laplacian operator, e is charge on the electron, Z is an atomic number, and r is the distance between particles.

Here, Ψ is a function of $3n$ coordinates where n is the total number of particles (nuclei and electrons), e.g., the x, y, and z Cartesian coordinates specific to each particle. If we work in Cartesian coordinates, the Laplacian has the form

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} . \tag{2.3}$$

The Hamiltonian operator in Eq. (2.2) is composed of kinetic and potential energy parts. The kinetic energy for a QM particle, however, is not expressed as $|\mathbf{p}|^2/2m$, but rather as the eigenvalue of the kinetic energy operator

$$T = -\frac{\hbar^2}{2m}\nabla^2. \tag{2.4}$$

According to Born interpretation, the square of an electronic wave function thus gives the electron density at any given point. If we integrate the probability of finding the particle over all space, then the result must be one as the particle must be somewhere:

$$\iiint |\Psi|^2 \, dxdydz = 1 . \tag{2.5}$$

Wave functions which satisfy this condition are said to be normalized. It is usual to require the solutions to the Schrödinger equation to be orthogonal:

$$\iiint \Psi_i^* \Psi_j \, dxdydz = 0 \;\; (i \neq j) \, . \tag{2.6}$$

Here, * denotes complex conjugation of Ψ_j.

A convenient way to express both the orthogonality of the different wave functions and the normalization conditions uses the Kröneckers delta:

$$\int \Psi_i \Psi_j \, d\tau = \delta_{ij} \, . \tag{2.7}$$

Here, $d\tau$ is $3n$-dimensional volume element.

Now, consider the result of taking Eq. (2.1) for a specific Ψ_i, multiplying $\Psi_j$on the left and integrating the resultant equation over τ.

This process gives

$$\int \Psi_j^* H \Psi_i \, d\tau = \int \Psi_j^* E_i \Psi_i \, d\tau \, . \tag{2.8}$$

Since the energy is a scalar value, we can write

$$\int \Psi_j H \Psi_i \, d\tau = E \delta_{ij} \, . \tag{2.9}$$

2.2. The Born-Oppenheimer approximation

Under typical physical conditions, the nuclei of molecular systems are moving far slower than the electrons. Recall the protons and neutrons are about 1800 times more massive than electrons and note the appearance of mass in the denominator of the kinetic energy terms of the Hamiltonian in Eq. (2.2).[1]

As such, it is convenient to decouple these two motions, and compute electronic energies for fixed nuclear positions. That is, the nuclear kinetic energy term is taken to be independent of the electrons, correlation in the attractive electron-nuclear potential energy term is eliminated, and the repulsive nuclear-nuclear potential energy term becomes a simply evaluated constant for a given geometry. Thus, the electronic Schrödinger equation is taken to be

$$(H_{\mathrm{el}} + \mathrm{V}_N)\Psi_{\mathrm{el}}(\mathrm{q}_i; \mathrm{q}_k) = E_{\mathrm{el}}\Psi_{\mathrm{el}}(\mathrm{q}_i; \mathrm{q}_k)\,. \qquad (2.10)$$

H_{el} includes only the first, third, and fourth terms on the right had side of Eq. (2.2), V_N is the nuclear-nuclear repulsion energy, and the electronic coordinates q_i are independent variables but the nuclear coordinates q_k are paprameters (and thus appear following a semicolon rather than comma in the variable list for Ψ).

The eigenvalue of the electronic Schrödinger equation is called the 'electronic energy'. Note that the term V_N is a constant for a given set of fixed nuclear coordinates.

Wave functions are invariant to the appearance of constant terms in the Hamiltonian, so in practice one almost always solves Eq. (2.10) without the inclusion of V_N, in which case the eigenvalue is sometimes called the 'pure electronic energy', and one then adds to this eigenvalue to obtain E_{el}.

2.3. Ionization energy

In 1933, the Dutch theoretical chemist T. Koopmans proposed that ionization energy is equal to the minus of the energy of the orbital (usually HOMO) from which the electron is ejected.[2-4]

The electronic energy of such a system can be calculated in a manner analogous to that for the hydrogen molecule. First, there is the energy of each electron moving in the field of the bare nuclei. For an electron in a molecular orbital χ_i, this contributes an energy H_{ii}^{core}.

If there are two electrons in the orbital then the energy is $2H_{ii}^{core}$ and for $N/2$ orbitals the total contribution to the energy will be[5]

$$\sum_{i=1}^{N/2} 2H_{ii}^{core} . \tag{2.11}$$

The interaction between each pair of orbitals ψ_i and ψ_j involves a total of four electrons if we consider the electron-electron terms. There are four ways in which two electrons in one orbital can interact in a Coulomb sense with two electrons in a second orbital, thus giving $4J_{ij}$. However, there are just two ways to obtain paired electrons from this arrangement, giving a total exchange contribution of $-2K_{ij}$.

Finally, the Coulomb interaction between each pair of electrons in the same orbital must be included; there is no exchange interaction because the electrons

have paired spins. The total electronic energy is thus given as:

$$E = 2\sum_{i=1}^{N/2} H_{ii}^{core} + \sum_{i=1}^{N/2}\sum_{j=i+1}^{N/2} (4J_{ij} - 2K_{ij}) + \sum_{i=1}^{N/2} 2J_{ij} \,. \tag{2.12}$$

A more concise form of this equation can be obtained if we recognize that $J_{ii} = K_{ii}$:

$$E = 2\sum_{i=1}^{N/2} H_{ii}^{core} + \sum_{i=1}^{N/2}\sum_{j=1}^{N/2} (2J_{ij} - K_{ij}) \,. \tag{2.13}$$

A Hartree-Fock calculation provides a set of orbital energies, ε_i. The energy of an electron in a spin orbital is calculated by adding the core interaction H_{ii}^{core} to the Coulomb and exchange interactions with the other electrons in the system:

$$\varepsilon_i = H_{ii}^{core} + \sum_{j=1}^{N/2} (2J_{ij} - K_{ij}) \,. \tag{2.14}$$

where, J_{ij} is often used to represent this Coulmb interation beween electrons in spin orbitals i and j. The energy due to exchange 'interaction' is often represented as K_{ij}. In a closed-shell system containing N electrons in $N/2$ orbials, there are two spin orbitals associated with each spatial orbital ψ_i : $\psi_i\alpha$ and $\psi_i\beta$.

One should be careful when applying Koopman's theorem and comparing the results with experimentally determined ionization energies. It is that the orbitals in the ionized state are assumed to be the same as in the unionized state; they are 'frozen'. This neglects the fact that the orbitals in the ionized state will be

different from those in the unionized state. The energy of the ionized state will thus tend to be higher than it should be, giving too large an ionization energies.[2,4]

2.4. Charge distribution

Charge density[3] at a point in an atom, molecule, or sold is the charge per unit volume at each location. For a one-electron system, it is equal to the absolute square of the wave function at the point of interest multiplied by the electronic charge, $-e$.

For a many-electron system, the total charge density is the sum of all the one-electron contributions. The product of the charge density at a location and the volume element $d\tau$ is equal to the charge inside the volume element.

In Cartesian coordinates, the expectation values of electric multipole moment operators are computed as[6]

$$\langle \mathbf{x}^k \mathbf{y}^l \mathbf{z}^m \rangle = \sum_i^{\text{atoms}} \mathrm{Z}_i x_i^k y_i^l z_i^m - \int \Psi^*(\boldsymbol{\tau}) \left(\sum_j^{\text{electrons}} x_j^k y_j^l z_j^m \right) \Psi(\boldsymbol{\tau}) d\boldsymbol{\tau} \quad , \qquad (2.15)$$

where the sum of $k, l,$ and m determines the type of moment (0 = monopole, 1 = dipole, 2 = quadrupole, etc.), Z_i is the nuclear charge on atom i, and the intergration variable $\boldsymbol{\tau}$ contains the x, y, and z coordinates of all of the electrons j.

The simplest moment to evaluate is the monopole moment, which has only the component $k = l = m = 0$, so that the operator becomes $\mathbf{1}$ and independent of coordinate system we have

$$\langle \mathbf{1} \rangle = \sum_{i}^{\text{atoms}} \mathrm{Z}_i - \sum_{j}^{\text{electrons}} \int \Psi_j(\boldsymbol{\tau}_j)\Psi_j(\boldsymbol{\tau}_j) d\boldsymbol{\tau}_j = \sum_{i}^{\textbf{atoms}} \mathrm{Z}_i - N \ , \qquad (2.16)$$

where N is the total number of electrons.

The monopole moment is thus the different between the sum of the nuclear charges and the number of electrons, i.e., it is the molecular charge. For the dipole moment, there are three possible components: x, y, or z depending on which of k, l, and m is one (with the others set equal to zero).

These are written as μ_x, μ_y, and μ_z. Experimentally, however, one rarely measures the separate components of the dipole moment, but rather the total magnitude, μ, which can be determined as

$$\langle \mu \rangle = \sqrt{\langle \mu_x \rangle^2 + \langle \mu_y \rangle^2 + \langle \mu_z \rangle^2} \ . \qquad (2.17)$$

The dipole moment measures the degree to which positive and negative charge are differential distributed relative to one another, i.e., overall molecular polarity.

2.5. References

[1] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), pp 96-98.

[2] Andrew R. Leach, *Molecular Modelling* (Anddison Wesley Longman Ltd, Harlow Essex, 1996), p. 110.

[3] P.W. Atkins, *Quanta* (Oxford University Press Inc., New York, 1991)

[4] Warren J. Hehre; Leo Radom.; Paul v. R. Schleer; John A. Pople, *Ab initio Molecular Orbital Theory* (John wiley & Sons Inc., New York, 1986), p. 24.

[5] Andrew R. Leach, *Molecular Modelling* (Anddison Wesley Longman Ltd, Harlow Essex, 1996), pp. 53-64.

[6] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), pp 275-276.

Chapter 3.

Computational Methods

3.1. Introduction

In this chapter, theoretical background, methodology, and computational method will be presented. In the first section, an introduction to theoretical background and methodology employed for our study will be indicated briefly.

In the second section, how the calculations were performed will be described in detail.

3.1.1. Density functional theory

The density functional theory (DFT) is widely applied to calculations of the electronic structure of atoms and molecules. Hohenberg and Kohn presented in 1964 that all the ground state properties of a system are functions of the charge density and their theorem is motivation of density functional theory.[1]

It means that an incorrect density gives energy above the true energy. Thus, it is most important to choose the successful candidate functional. The reasonable density will determine suitable wave function and Hamiltonian. That being the case, we can evaluate the energy expectation value

$$\langle \Psi_{\text{candidate}} | H_{\text{candidate}} | \Psi_{\text{candidate}} \rangle = E_{\text{candidate}} \geqq 0 \,. \qquad (3.1)$$

This expression indicates that the expectation value by the variational principle, must be greater than or equal to the true ground state energy.[2] So, in principle, we can select more optimized densities to be closer and correct.

According to the theorem, the total electronic energy should be expressed as a function of the electron density ρ:

$$E(\rho) = E_{KE}(\rho) + E_C(\rho) + E_H(\rho) + E_{XC}(\rho) \ , \qquad (3.2)$$

where $E_{KE}(\rho)$ is the kinetic energy, $E_C(\rho)$ is the electron-nuclear interaction term, $E_H(\rho)$ is the electron-electron Coulombic energy, and $E_{XC}(\rho)$ contains the exchange and correlation contributions.[1]

All of the electron-electron interactions are thus contained within the $E_H(\rho)$ and $E_{XC}(\rho)$ terms. To perform a density functional calculation, it is necessary to rewrite equation (3.2) in terms of the density and then the optimized energy with respect the density, subject to any constraints on the system. The difficulty comes from the electron-electron interaction term in the correct Hamiltonian.[2]

For the exchange-correlation term $E_{XC}[\rho(\mathbf{r})]$, some approximation must be made. The most common way to obtain this contribution uses the so-called local density approximation (LDA), which is based upon a model called the uniform electron gas. In a uniform electron gas, the electron density is constant throughout all space.

The exchange-correlation energy can be determined for this model. In the commonly used Kohn-Sham implementation, the density is represented as if it were derived from a single Slater determinant with orthonormal orbitals.[3]

The various components that contribute to the energy (3.2) must first be expressed in terms of the density. The charge density at a point $\mathbf{r}$ can be written as the sum over occupied molecular orbitals of ψ^2:

$$\rho(\mathbf{r}) = \sum_{\mathrm{i}=1}^{\mathrm{occ}} |\psi_{\mathrm{i}}(\mathbf{r})|^2 \,. \tag{3.3}$$

In ref 3, the use of these 'Kohn-Sham orbitals' with equation (3.3) enables the energy to be optimized by solving a set of one-electron Schrödinger equations as follow

$$\left[-\frac{\nabla^2}{2} + r_{\mathrm{nuclear}}(\mathbf{r}) + \int d\mathbf{r}' \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} + r_{\mathrm{XC}}(\mathbf{r})\right]\psi_{\mathrm{i}}(\mathbf{r}) = \varepsilon_{\mathrm{i}}\psi_{\mathrm{i}}(\mathbf{r}) \tag{3.4}$$

and exchange-correlation functional is the derivative of the exchange-correlation energy with respect to the density:

$$r_{\mathrm{XC}}(\mathbf{r}) = \frac{\delta \mathrm{E}_{\mathrm{XC}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \,. \tag{3.5}$$

The energy is given by

$$E(\rho) = \sum_{\mathrm{i}=1}^{N} \int \psi_{\mathrm{i}} \left(-\frac{\nabla^2}{2}\right) \psi_{\mathrm{i}} d\mathbf{r} + \int r_{\mathrm{nuclear}} \rho(\mathbf{r}) d\mathbf{r}$$

$$+ \frac{1}{2} \iint d\mathbf{r}\, d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \mathrm{E}_{\mathrm{XC}}[\rho(\mathbf{r})] \,. \tag{3.6}$$

DFT calculations requires about the same amount of computation resource as Hatree-Fock theory but they include the effect of electron correlation.[4] This is one of the key advantages of the density functional approach.

3.1.2. Density functionals

(1) B3LYP

Stevens **et al.** modified B3PW91 and developed this method in 1994.[5] The name of functional, B3LYP, implies its use of a three-parameter scheme, as well as the generalized gradient approximation (GGA) exchange and correlation functional B and LYP, respectively.

Let us recall the energy for Hartree-Fock (HF) and density functional theory (DFT). They can be expressed in terms of simplified formula[6] for convenience in equations (3.7) and (3.8), respectively.

$$E_{HF} = \langle h\rho \rangle + V + \frac{1}{2}\langle PJ(\rho)\rangle + E_X[\rho]\,. \qquad (3.7)$$

Here, $\langle h\rho \rangle$ is the one-electron (kinetic + potential) energy, V is the nuclear repulsion energy, $\frac{1}{2}\langle PJ(\rho)\rangle$ is the classical coulomb repulsion of the electrons, and $E_X[\rho]$ is the exchange energy resulting from the quantum nature of electrons.

In the Kohn-Sham formulation of density functional theory, energy E_{KS} is given as,

$$E_{KS} = \langle h\rho \rangle + V + \frac{1}{2}\langle PJ(\rho)\rangle + E_X[\rho] + E_C[\rho]\,. \qquad (3.8)$$

In equation (3.8), $E_C[\rho]$ is the correlation functional. The exact exchange (HF) for a single determinant is replaced by a more general expression, that is, the

exchange-correlation functional, which can include terms accounting for both the exchange and the electron correlation energies, the latter not being present in Hartree-Fock theory as above, i.e., HF theory also can be regarded as a special case of DFT, in which $\mathrm{E_X}[\rho]$ is given by the exchange integral $-\frac{1}{2}\langle \mathrm{PK}(\rho)\rangle$ and $\mathrm{E_C}[\rho] = 0$ within the Kohn-Sham formulation.

A "hybrid" exchange-correlation functional is usually constructed as a linear combination of the HF exact exchange functional, $\mathrm{E_x^{HF}}$ given as,

$$\mathrm{E_X^{HF}} = \frac{1}{2}\sum_{i,j}\iint \psi_i^*(\mathbf{r_1})\psi_j^*(\mathbf{r_1})\frac{1}{\mathbf{r_{12}}}\psi_i(\mathbf{r_2})\psi_j(\mathbf{r_2})d\mathbf{r_1}d\mathbf{r_2}\,, \qquad (3.9)$$

and the parameters determining the weight of each individual functional are typically specified by fitting the functionals predictions to experimental or accurately calculated thermochemical data.

B3LYP is a kind of hybrid density functional theory models to improve calculation accuracy by using correction to exchange-correlation energy and "hybrid" means a mixing other exchange-correlation energy functional in density functional theory (DFT) such as equation (3.10).[5,7-9]

Note the conventional definition of $\mathrm{E_{XC}}$ is equal to sum of exchange energy and correlation energy.

$$E_{\mathrm{XC}}^{hyb} = (1-a)\mathrm{E}_{\mathrm{XC}}^{\mathrm{DFT}} + a\mathrm{E}_{\mathrm{X}}^{\mathrm{HF}} = \mathrm{E}_{\mathrm{XC}}^{\mathrm{DFT}} + a\left(\mathrm{E}_{\mathrm{X}}^{\mathrm{HF}} - \mathrm{E}_{\mathrm{X}}^{\mathrm{DFT}}\right) - a\mathrm{E}_{\mathrm{C}}^{\mathrm{DFT}}, \qquad (3.10)$$

where a is an empirical constant to be optimized by appropriate fits to firm experimental data.

Exchange-correlation functional for B3LYP is

$$E_{\mathrm{XC}}^{\mathrm{B3LYP}} = E_{\mathrm{XC}}^{\mathrm{LSDA}} + \alpha_0\left(E_{\mathrm{X}}^{\mathrm{exact}} - E_{\mathrm{X}}^{\mathrm{LSDA}}\right) + \alpha_{\mathrm{x}}\Delta\mathrm{E}_{\mathrm{X}}^{\mathrm{B}} + \alpha_{\mathrm{c}}\Delta\mathrm{E}_{\mathrm{c}}^{\mathrm{LYP}}, \qquad (3.11)$$

where α_0, α_{x}, and α_{c} are *semiempirical coefficients* to be determined by an appropriate fit to experimental data, $E_{\mathrm{X}}^{\mathrm{exact}}$ is extract exchange energy, $\Delta\mathrm{E}_{\mathrm{X}}^{\mathrm{B}}$ is Becke's gradient correction (to the LSDA) for exchange, and $\Delta\mathrm{E}_{\mathrm{c}}^{\mathrm{LYP}}$is gradient correction for correlation if LYP.[5,7]

Although density functional theory is widely used in the field of computational chemistry, the most popular density functional, B3LYP, has some serious shortcomings:[10] (1) it results in irrelevant for transition metals, (2) it systematically underestimates reaction barrier heights; (3) it is inaccurate for interactions dominated by medium range correlation energy, such as van der Waals attraction, $\pi\cdots\pi$ stacking, and alkane isomerization energies.

The shortcomings originate from the parameter fitting carried out near the region of equilibrium point.

(2) M05-2X

Density functionals using the local spin density approximation (LSDA) depends only on density, however, ones in the generalized gradient approximation (GGA) depends on density and its reduced gradient.[11] GGA functionals have been shown to give more accurate predictions for thermochemistry than LSDA ones, but they still underestimate barrier heights.

Two additional variables, the spin kinetic energy densities are included in the functional form; such functionals are called meta-GGAs. LSDAs, GGAs, and meta-GGAs are "local" functionals because the electronic energy density at a single spatial point depends only on the behavior of the electronic density and kinetic energy.[10, 11]

The long-range correction scheme is inapplicable to most of generalized gradient approximation (GGA) exchange functions. This is because conventional functionals are not usually derived from the corresponding first-order density matrix, although it is necessary to define the dependence of an exchange functional on electron-electron distance.[12]

M05-2X[10,11,13] developed by D. G. Truhlar's group was designed to be suitable for main-group thermochemical kinetics, and noncovalent interactions (weak interaction such as hydrogen bonding, $\pi \cdots \pi$ stacking, interactions energies of nucleobases) by employing a high-nonlocality functional with double the amount of nonlocal exchange (2X).

It adopts hybrid meta-GGA exchange and correlation functionals including all include terms that depend on the kinetic energy density. M05-2X of M05 series was reported to have best performance for weak interactions excluding metal. Showing the detail of the functional, it is composed of four parts: electron spin density, density gradient, kinetic energy density, and Hartree-Fock (HF) exchange.

The Hybrid exchange-correlation energy can be written as follows:[13]

$$E_{\mathrm{XC}}^{\mathrm{hyb}} = E_{\mathrm{X}}^{\mathrm{HF}} + \left(1 - \frac{X}{100}\right)\left(E_{\mathrm{X}}^{\mathrm{DFT}} - E_{\mathrm{X}}^{\mathrm{HF}}\right) + E_{\mathrm{C}}^{\mathrm{DFT}} \quad , \qquad (3.12)$$

where $E_{\mathrm{X}}^{\mathrm{HF}}$ is the nonlocal HF exchange energy, X is the percentage of HF exchange in the hybrid functional, $E_{\mathrm{X}}^{\mathrm{DFT}}$is the local DFT exchange energy, and $E_{\mathrm{C}}^{\mathrm{DFT}}$ is the local DFT correlation energy. In equation (3.12), it is very important that nonlocal HF exchange, $E_{\mathrm{X}}^{\mathrm{HF}}$, be used in M05-2X. On the other hand, that of B3LYP results in density functional theory.

One can see that the total correlation energy for a DFT calculation is modeled as the sum of the dynamic correlation energy given by $E_{\mathrm{C}}^{\mathrm{DFT}}$ and the nondynamic correlation energy contained in

$$\left(1 - \frac{X}{100}\right)\left(E_{\mathrm{X}}^{\mathrm{DFT}} - E_{\mathrm{X}}^{\mathrm{HF}}\right).$$

The M05-2X functional is classified hybrid meta-generalized gradient approximations (hybrid meta-GGAs) because all components except HF exchange are local properties of the density.

Especially, it was reported that the M05-2X give the best performance for noncovalent interactions.[10,13]

3.2. Computational details

All calculations in this thesis were carried out using Gaussian 09 Rev. A01 software package[06] to employ the M05-2X method.[11] The neutral and cationic geometries of conformers of the three aromatic amino acids in the ground state were optimized at M05-2X/6-311+G* level and the same calculation was also carried out with B3LYP/6-311+G* in order to compare with the M05-2X results.

Relative energies between conformers belong to a same aromatic amino acid also were evaluated through the geometry optimization. The optimized structure in corresponding transition state (TS) was obtained for the most probable reaction path between two neutral conformers belonging to the same subgroup, which was confirmed using IRC profiles[06]. Ionization energies for the three aromatic amino acids were also computed with same method by setting the charge and the spin multiplicity to +1 and 2, respectively.

In order to get more comparable results between theoretical estimates and experimental data for all the three aromatic amino acids, zoro-point energy correction[06] was performed together with M05-2X/6-311+G*. Naturally, charge distribution for cations also was obtained using NBO[06] analysis.

The expectation values of the square of the total electron spin, $<S^2>$ was examined in order to check the spin contamination of cationic wave functions. The values of $<S^2>$ calculated by unrestricted M05-2X (B3LYP) were

approximately 0.750 (0.750) for all conformers of cationic three aromatic amino acids. We used Gaussian 09 Rev. A01 installed in the Prof. Hayashi's cluster machine of Taiwan National University. The GaussView 5 (Gaussian Inc.)[15] was employed in order to extract their initial Cartesian coordinates and visualize the result after normal termination for conformers of the three aromatic amino acids, respectively.

Figure 3.1 shows an example of the input file for the Gaussian 09 job and it consists of command lines and information of geometry of molecule[4]. The symbol, % represents command line in the input file. The first command "Nprocshared=" designates the number of processors to use for the Gaussian 09 job. The second one, "Mem=" allocates memory resource for the calculation. The third one, "chk=" saves the calculation result into the predetermined checkpoint file in the command line. The forth command line includes calculation method (M052X), basis set (6-311+G*), calculation purpose (Opt, geometry optimization for the molecule) and so on.

The next section in the input file specifies the charge and the spin multiplicity (0 1), Cartesian coordinate for the molecule. Note insertion of a blank in the last line.

```
%Nprocshared=4
%Mem=8GB
%chk=M052X_1-X.chk
#p Iop(7/8=11) M052X/6-311+G* Opt(Calcall, Maxcyc=300) Pop=NBO

M052X_1-X

 0  1
 C              0.77112500          0.26916200         -0.68296800
 C              0.97200800         -1.09156500         -0.44121800
 C              2.13060700         -1.53137100          0.18628600
 C              3.10559600         -0.62155200          0.58044300
 C              2.91861100          0.73345500          0.34003500
 C              1.75785500          1.17257000         -0.28809900
 H              0.21938800         -1.80376100         -0.75218400
 H              2.27380800         -2.58856000          0.36408600
 H              4.00761100         -0.96788400          1.06617100
 H              3.67616200          1.44740400          0.63403800
 H              1.62172900          2.22908800         -0.48887500
 C             -0.50534100          0.76267600         -1.31615300
 C             -1.65718500          0.85416100         -0.30031900
 H             -0.81784600          0.09781000         -2.11931700
 H             -0.34381000          1.75408500         -1.74118600
```

Figure 3.1. An example of the input file for the Gaussian 09 job.

3.3. References

[1] Andrew R. Leach, *Molecular Modelling* (Anddison Wesley Longman Ltd, Harlow Essex, 1996), p.p. 528-530.

[2] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), pp 233-242.

[3] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

[4] J. B. Foresman and M. J. Frisch, *Exploring Chemistry with Electronic Structure Methods*, 2nd ed. (Gaussian, Inc., Pittsburgh, 1996)

[5] Christopher J. Cramer, *Essential of computational chemistry* (John Wiley & Sons, Ltd., West Sussex, 2002), pp 248-251.

[6] Gaussian 09, Revision A.1, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.;

Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2009**.

[7] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[8] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

[9] Lee, C.; Yang, W.; Parr, R.G. *Phys. Rev. B* **1988**, *37*, 785.

[10] Zhao, Y.; Truhlar, D. G. *Accounts Chem. Res.* **2008**, *41,* 157.

[11] Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109,* 5656.

[12] Iikura, H.; Tsuneda, T.; Yanai, T.; and Hirao, K. *J. Chem. Phys.* **2001**,115, 3540.

[13] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[14] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583.

[15] GaussView, Version 5, Dennington, Roy; Keith, Todd; Millam, John. Semichem Inc., Shawnee Mission, KS, **2009**.

Chapter 4.

Structure and Energy for the Three Aromatic Amino Acids in Neutral

4.1. Introduction

Assignment of structures between their conformers is not easy as one think in spite of remarkable development of experimental techniques and computational methods. Recently, the trend of preference for calculation method is going to be divided depending on study of purpose in chemistry and physics, that is, the most popular hybrid density functional theory (DFT) like B3LYP can be the best choice if research taget is thermochemistry but it had better to avoid the method if target of study involves weak interaction. [1]

For L-phenylalanine, two conformers of the observed six conformers finally became reassigned and some of the stable conformers predicted by B3LYP calculation could not be observed experimentally.[2-4]

So far, most of the theoretical analyses of conformer spectra have been conducted by using B3LYP excluded long-range effect even though it was certainly known that the three aromatic amino acids (L-phenylalanine, L-tyrosine, and L-tryptophan) are noncovalent long-range interactions such as intramolecular hydrogen bonding.[6-11]

Therefore, it is worth applying a new DFT functional M05-2X to neutrals and cationic conformers of these aromatic amino acids in order to theoretically confirm their assignments and clarify correspondence between observed and predicted conformers. In particular, unlike L-phenylalanine, the other cationic

conformers corresponding to L-tyrosine and L-tryptophan remaining largely unexplored have a value to be investigated. In the next section, theoretical methods and procedures are briefly described.

Section 4.3 provides explicit structural assignment of neutral and cationic conformers of all the three aromatic amino acids based on the M05-2X results of optimized geometrical structure which is a kind of conformation-dependent properties.[12,13] Assignment of some conformers for L-phenylalanine have been argued so far.

Therefore, the most probable reaction path of the argued conformers is theoretically investigated along with intrinsic reaction coordinate (IRC)[14] profile in section 4.3. The results Their optimized structures and energies obtained by B3LYP are also shown in order to compare with that of M05-2X.

4.2. Computational procedure

The each geometry of neutral conformers of L-phenylalanine in the ground state is optimized at the M05-2X/6-311+G* level including zero-point energy correction (ZPE_{corr}) with scanning for five dihedral angles rotation by setting the charge and spin multiplicity to 0 and 1, respectively: C_γ–O = 0°, 180°; C_π–C_β = 60°, 120°,180°; C_β –C_α = -120°, -60°, 0°, 60°, 120°, 180°; C_α–N = -120°, -60°, 0°, 60°, 120°, 180°; C_α–C_γ = 0°, 30°, 90°, 150°, 210°, 270°, 330° in Figure 4.1(a).

A total of 1296 initial geometries of the multiconformer exist, and they are converged to 39 stable local minima and 38 local minima in B3LYP with 6-311+G* and the results obtained from B3LYP reproduce exactly the report by Lin **et al.**[15]

On the other hand, that of cationic conformers of L-phenylalanine in ground state are also optimized at the M05-2X/6-311+G* level including zero-point energy correction (ZPE_{corr}) by setting the charge and spin multiplicity to +1 and 2 using their optimized neutral structures as a initial geometry, respectively.

For L-phenylalanine, the optimized structure in the corresponding transition state (TS) is obtained for the most probable reaction path between two neutral conformers belonging to the same subgroup, which is confirmed using IRC profiles.[14]

In the same manner, the geometries of neutral conformers of L-tyrosine and L-tryptophan in the ground state are optimized at the M05-2X level with the 6-311+G* basis set, based on the similar conformations of the backbone (alanine) in L-phenylalanine.[16]

The eighteen most stable neutral conformers are represented for both L-tyrosine and L-tryptophan by considering the orientation of the phenolic –OH group for L-tyrosine and the asymmetry of the indole ring for L-tryptophan.

In similar of L-phenylalanine, for each conformer of L-tyrosine and L-tryptophan radical cation in the ground state, geometry optimization is calculated at M05-2X level with the 6-311+G* basis set by setting the charge and the spin multiplicity to +1 and 2, respectively.

4.3. Results and discussion

4.3.1. Geometrical structures of neutral, and cationic conformers for the three aromatic amino acids

(1) L-phenylalanine

Figure 4.1a(b) show the nine stable neutral L-phenylalanine conformers and six cationic ones in their ground states. Basically, they were calculated by using M05-2X optimization and conformers obtained by using the popular (DFT) B3LYP, which were different from the M05-2X results, are illustrated together.

A comparison of M05-2X and B3LYP results on their molecular constants for L-phenylalanine conformers is shown in Table 4.1. Here, dihedral angles, N–C_α–C_β–C_π and C_α–NH_2, and bond lengths, C_α–C_β, C_α–N, and C_β–C_π, which characterize each conformer, are listed.

The two functionals, M05-2X and B3LYP, gave almost the same results regarding geometrical structures for other groups such as the phenyl ring and the carboxyl groups in both the neutrals and the cations.

The neutral conformers of six alphabetic capital letters, A, B, C, D, E, and X, in Figure 4.1a(b) are quoted from report by Snoek **et al.**, in which their assignment was carried out based on a combination of UV-UV and IR-UV double-resonance

spectroscopy of jet-cooled L-phenylalanine and coupled with *ab* initio computations.[5] The nine Roman numerals I, II, ..., VIII, and IX are numbered in order of relative energies calculated using the MP2/6-311G** level of theory.

Neutral conformers IV, V, and VIII, which were predicted by the B3LYP results, have not been experimentally identified. The M05-2X results indicate that neutral conformer V belongs to subgroup II while cationic conformer V belongs to subgroup I.

This is totally different from the results obtained by B3LYP: both the neutral and the cationic conformers of V were assigned to subgroup II. Such a hybrid type of conformer that has different characters between the neutral and caion is one of the newly found results using the new DFT functional that takes into account noncovalent interactions as shown in Figure 4.1(b).

First, the conformer IV was assigned as observed structure E in ref 1, but IX was assigned later because of the obvious evidence in photo ionization energy (PIE) curve, shown by Lee **et al.**[7] According to ref 7, observed conformer E should belong to subgroup II based on the value of ionization energy whereas conformer IV belong to subgroup I, and instead conformer IX belong to subgroup II from the M05-2X results.

In previous papers on geometrical structures obtained by using the B3LYP method there was no clear difference in ionization energies or vibrational frequencies in the neutral ground state between V and VII conformers.[5,8,9]

Furthermore, there were no deviations of cationic structures, that is, the two conformers commonly showed elongation of the C_α–C_β bond and planarity of C_α–NH_2. (see Table 4.1) There has been heated debate about whether observed conformer A was to be assigned to conformer V or VII.[2,5]

Experimentally, conformer V was assigned to conformer A by Snoek **et al.**[5] Later, it was rationalized from UV rotational band contour analysis by ref 2 that conformer A corresponds not to conformer V but to conformer VII.

The results obtained by using the M05-2X method, in contrast, show no elongation of the C_α–C_β bond and a nonplanar, pyramidal structure for C_α–NH_2 on conformer V in Table 4.1.

Table 4.2 shows relative energies including zero-point energy correction (ZPE_{corr}) of the conformers, which were calculated by both methods. The original Roman numerals were used to avoid confusion even though their energy ordering is different from that of M05-2X in Table 4.2.

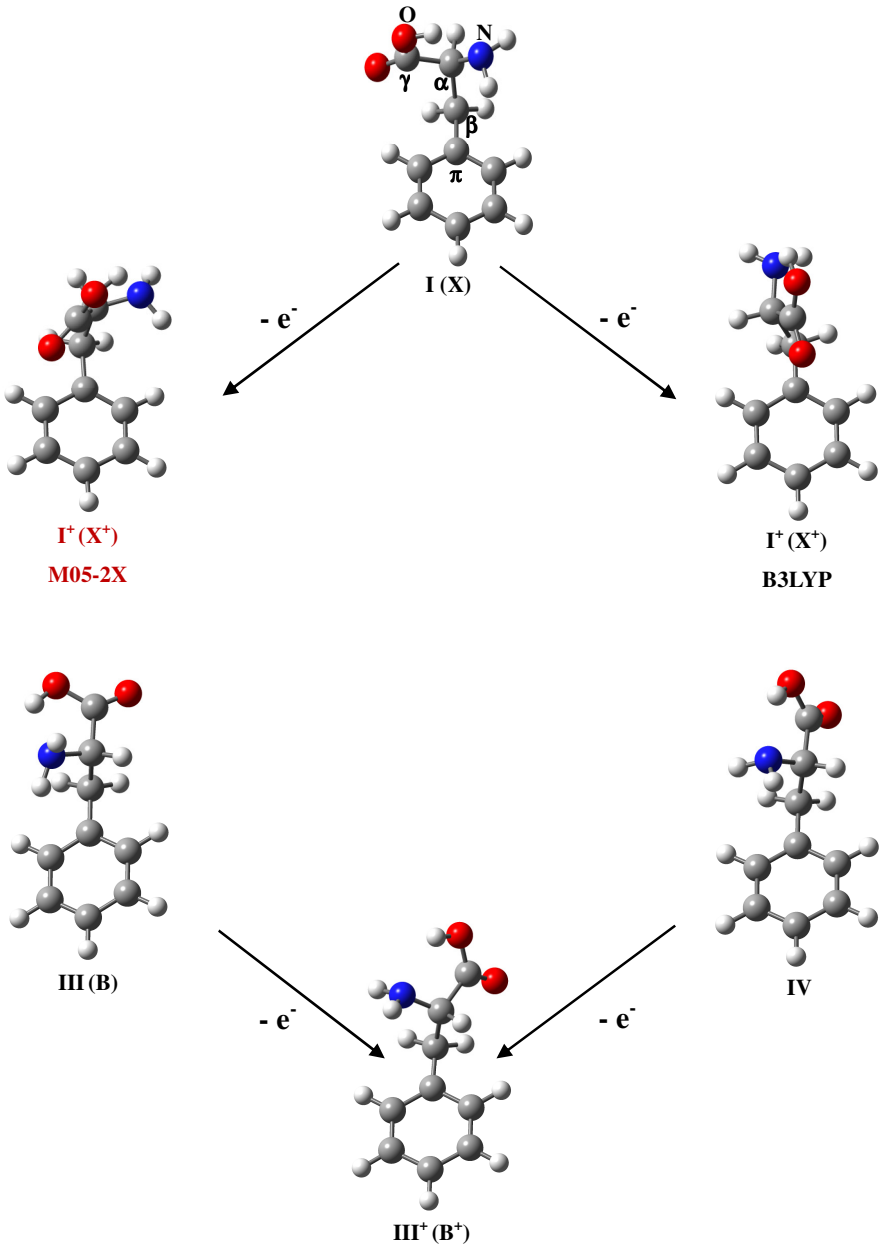


Figure 4.1(a) Three stable structures of conformers belonging to subgroup I in neutrals (cations) for L-phenylalanine.

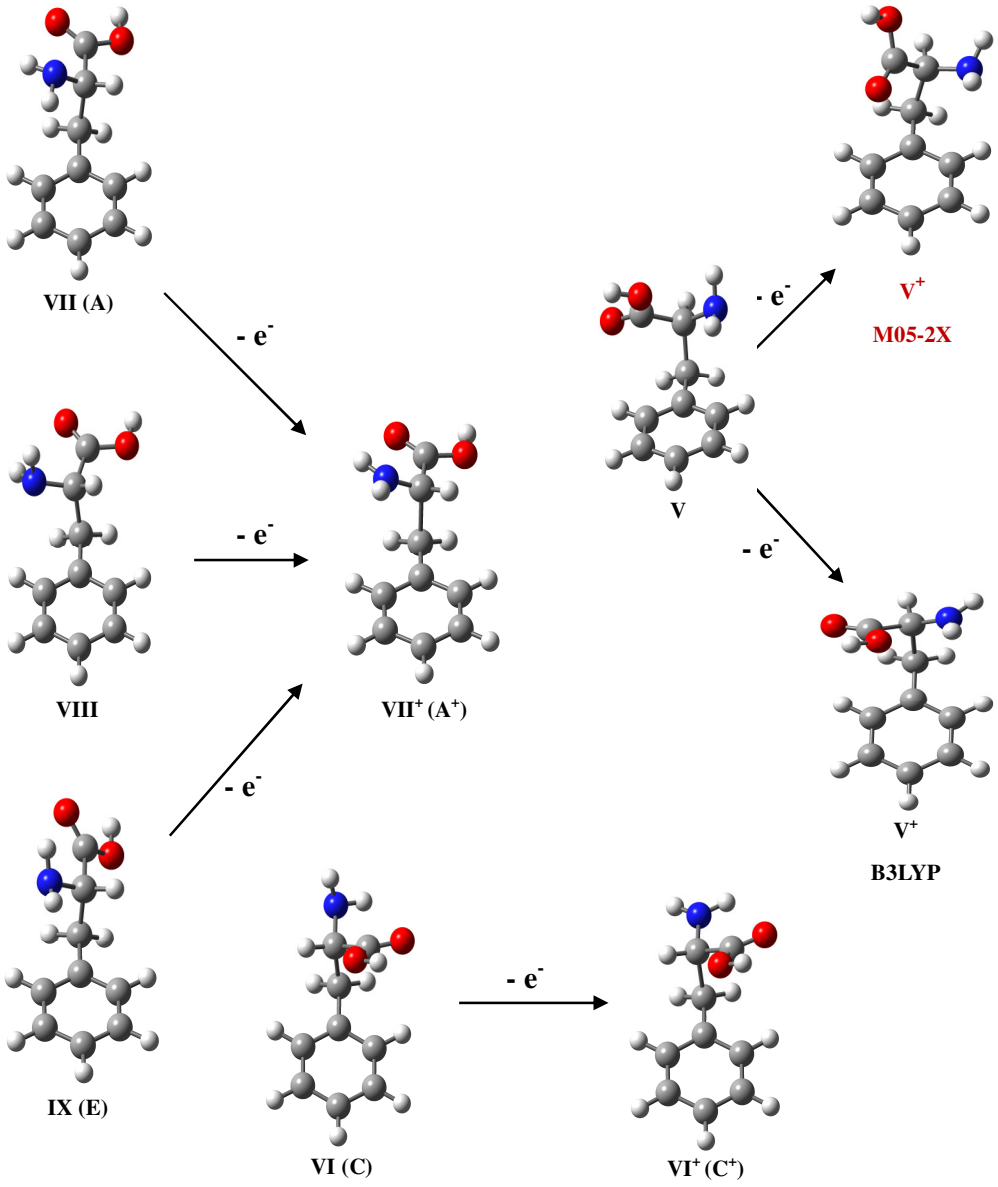


Figure 4.1(b) Six stable structures of conformers belonging to subgroup II in neutrals (cations) for L-phenylalanine.

Table 4.1. Calculated molecular constants for nine conformers of L-phenylalanine in neutrals and those for cations in parentheses.

(a) M05-2X, neutral (cation)

Sub.	Experiment Assignment	Conf.	Dihedral angle(°)		Bond length(Å)		
			C_π-C_β-C_α-N	C_α–NH_2	C_α–C_β	C_α–N	C_β– C_π
I	X	I	52(**88**)	123(131)	1.54(1.57)	1.46(1.44)	1.51(1.49)
I	B	III	-63(-86)	122(137)	1.54(1.59)	1.46(1.42)	1.51(1.46)
I		IV	-60(-86)	121(137)	1.53(1.59)	1.46(1.42)	1.51(1.46)
II (I)		**V**	62(64)	**120(127)**	**1.54(1.56)**	1.45(1.44)	1.51(1.48)
II	A	VII	-60(-69)	122(176)	1.55(1.77)	1.45(1.36)	1.50(1.45)
II	C	VI	178(171)	120(168)	1.54(1.79)	1.45(1.36)	1.51(1.44)
II	D	II	62(73)	119(178)	1.54(1.78)	1.45(1.36)	1.51(1.44)
II	E	IX	-62(-69)	123(176)	1.54(1.77)	1.45(1.36)	1.51(1.44)
II		VIII	-73(-69)	120(176)	1.54(1.77)	1.45(1.36)	1.51(1.45)

(b) B3LYP, neutral (cation)

Sub.	Experiment Assignment	Conf.	Dihedral angle(°)		Bond length(Å)		
			C_π-C_β-C_α-N	C_α–NH_2	C_α–C_β	C_α–N	C_β– C_π
I	X	I	52(**172**)	123(133)	1.55(1.56)	1.47(1.44)	1.51(1.50)
I	B	III	-64(-84)	122(143)	1.55(1.58)	1.47(1.42)	1.51(1.48)
I		IV	-64(-84)	122(143)	1.54(1.58)	1.47(1.42)	1.51(1.48)
II		**V**	61(72)	**121(179)**	**1.55(1.70)**	1.45(1.39)	1.51(1.47)
II	A	VII	-61(-72)	122(176)	1.57(1.70)	1.45(1.39)	1.51(1.47)
II	C	VI	-175(177)	121(165)	1.54(1.70)	1.46(1.39)	1.51(1.47)
II	D	II	62(73)	120(178)	1.55(1.70)	1.45(1.39)	1.51(1.46)
II	E	IX	-63(-73)	123(176)	1.55(1.70)	1.46(1.39)	1.51(1.46)
II		VIII	-73(-72)	120(176)	1.56(1.70)	1.45(1.39)	1.51(1.47)

Table 4.2. Relative energies in kcal/mol of the nine stable conformers in the neutral (cationic) ground state of L-phenylalanine.

Sub.	Experimental Assignment	Conf.	Neutral		Cation	
			M05-2X	B3LYP	M05-2X	B3LYP
I	X	I	0.00	0.11	6.29	1.68
I	B	III	0.62	0.00	10.03	7.29
I		IV	0.79	0.17	10.03	7.29
II		V	1.15	1.64	1.46	0.96
II	A	VII	1.50	1.00	0.71	0.00
II	C	VI	1.56	1.08	0.17	0.04
II	D	II	0.47	0.87	0.00	0.22
II	E	IX	1.71	1.22	0.71	0.00
II		VIII	2.06	1.09	0.71	0.00

(2) L-tyrosine

For L-tyrosine, eight band origins (4–7, 14–17) were observed by fluorescence-detected UV-UV and IR-UV double-resonance spectroscopy.[19] Figure 4.2a(b) show only the eight valid neutral conformers of L-tyrosine and their corresponding cations among the 18 optimized structures obtained by M05-2X: (a) for subgroup I, (b) for subgroup II.

Particularly, the cationic structures belonging to subgroup II obtained by the popular B3LYP are different from those from M05-2X and shown in parentheses.

A comparison of M05-2X and B3LYP calculations is shown in Table 4.3, which lists the dihedral angles ($N–C_\alpha–C_\beta–C_{\pi1}$, $C_\alpha–NH_2$, $N–C_\alpha–C_\gamma–O_1$) and the bond lengths ($C_\alpha–C_\beta$, $C_\beta–C_{\pi1}$, $C_\alpha–N$, and $C_{\pi2}–O_2$).

The two functionals, M05-2X and B3LYP, yielded nearly the same results with regard to the neutral structure, but the radical cations of subgroup II show different conformational structures depending on the functional used. This is mainly due to the effect of long-range noncovalent interactions, which is properly accounted for in the M05-2X functional but not in the B3LYP results.[21–24]

It should be noted in Table 4.3 that L-tyrosine conformers of subgroup II show no drastic change in geometry upon ionization, such as $C_\alpha–C_\beta$ elongation and planarity of the $C_\alpha–NH_2$ as seen in the subgroup II conformers of L-phenylalanine.[7,16] The neutral species shown in Figure 4.2a(b) are identical to

those reported by Ebata and co-workers.[19] The subgroup classification for L-tyrosine is based on the directionality of the intramolecular hydrogen bonding, as in the case of L-phenylalanine: $-COOH \rightarrow -NH_2$ for subgroup I and $-NH_2 \rightarrow -OCOH$ for subgroup II.

There are two different, nearly isoenergic, orientations of the phenolic –OH group, which are denoted by subscript r (right) or l (left). Roman numerals (I, II, III, and VII) were adopted from L-phenylalanine labels[16] because of the similarity between L-tyrosine and L-phenylalanine, with the –OH group in L-tyrosine being the major difference.

Ebata and co-workers observed eight origin bands (4, 5, 6, 7, 14, 15, 16, and 17) in their fluorescence excitation spectra[19] and proposed that pairs of 4 and 6, 5 and 7, and 16, and 17 correspond to two rotational isomers arising from the different orientations of phenolic –OH.

Table 4.4 shows relative energies of L-tyrosine conformers in neutrals (cations), which were calculated by using M05-2X and conventional B3LYP methods. Here, zero point energy corrections (ZPE_{corr}) were taken into account. According to the results in Table 4.4, these sets of rotational conformers are no different energetically.

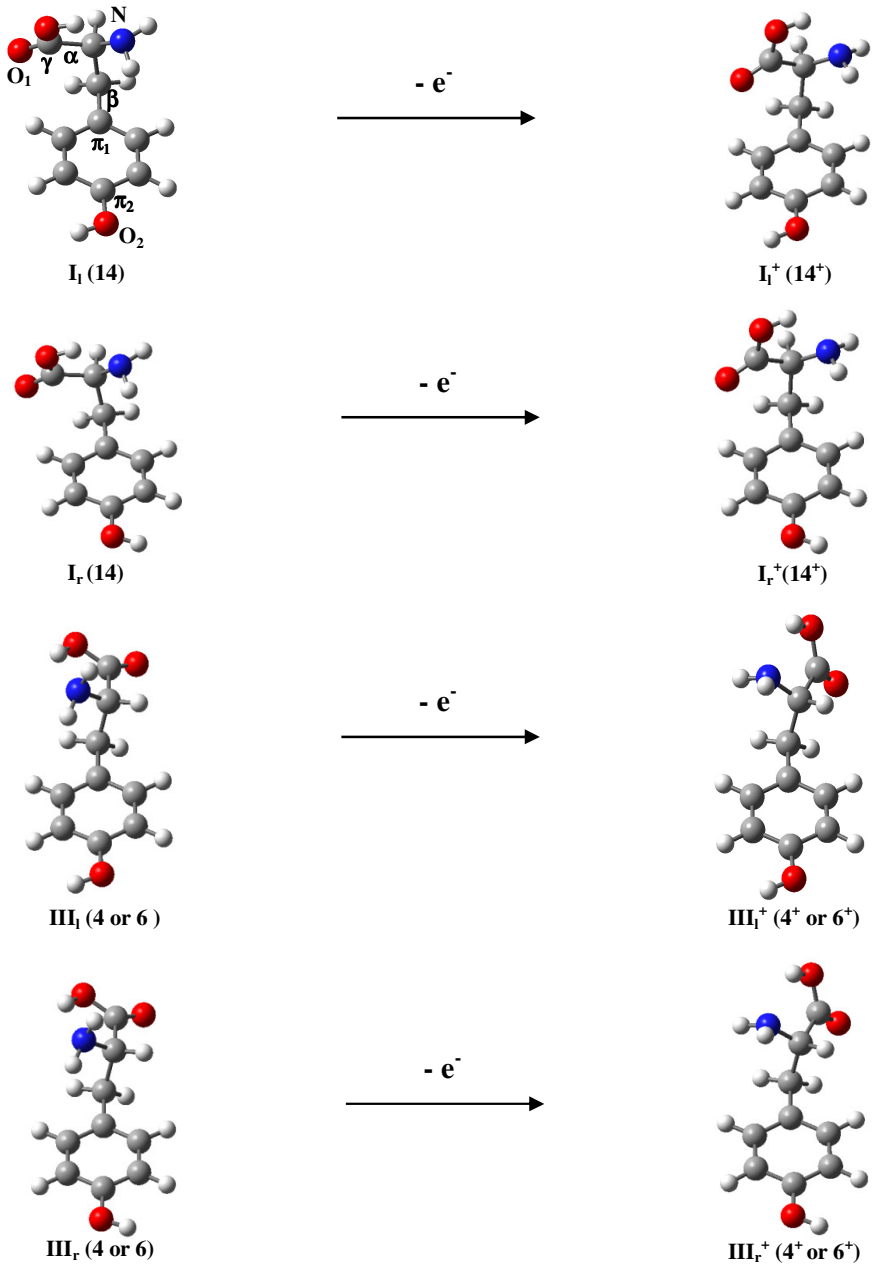


Figure 4.2(a) The four stable structures of representative L-tyrosine conformers belonging to subgroup I in neutrals (cations).

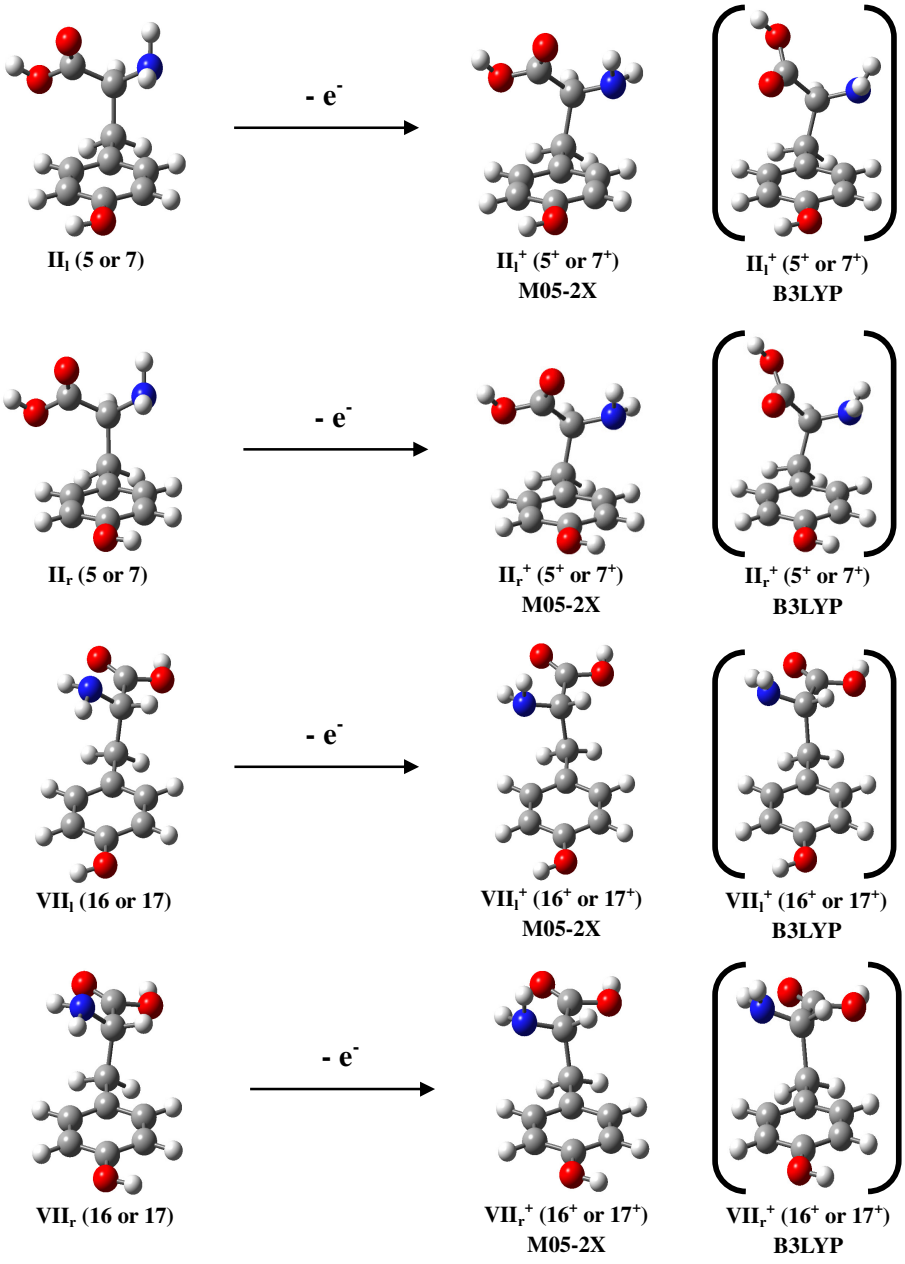


Figure 4.2(b) The four stable structures of representative L-tyrosine conformers belonging to subgroup II in neutrals (cations).

Table 4.3. Molecular constants for eight L-tyrosine conformers in neutrals and those for cations in parentheses.

Sub.	Origin Bands	Conf.	Dihedral angle (°)			Bond length (Å)			
			$C_{\pi 1}$-C_β-C_α-N	C_α-NH_2	O_1-C_γ-C_α-N	C_α-C_β	C_β-$C_{\pi 1}$	C_α-N	$C_{\pi 2}$-O_2
(a) M05-2X, neutral (cation)									
I	14	I_l	52 (70)	123 (132)	-169 (-150)	1.54 (1.56)	1.51 (1.48)	1.46 (1.44)	1.36 (1.31)
I	14	I_r	51 (68)	123 (132)	-169 (-151)	1.54 (1.56)	1.51 (1.48)	1.46 (1.44)	1.36 (1.31)
I	4 or 6	III_l	-63 (-80)	122 (132)	163 (-161)	1.54 (1.56)	1.51 (1.47)	1.46 (1.44)	1.36 (1.31)
I	4 or 6	III_r	-63 (-82)	122 (132)	163 (-162)	1.54 (1.56)	1.51 (1.47)	1.46 (1.44)	1.36 (1.31)
II	5 or 7	**II_l**	62 (66)	119 (125)	-1 **(0)**	1.54 (1.57)	1.51 (1.48)	1.45 (1.44)	1.36 (1.31)
II	5 or 7	**II_r**	62 (67)	119 (124)	0 **(0)**	1.54 (1.57)	1.51 (1.48)	1.45 (1.44)	1.36 (1.31)
II	16 or 17	**VII_l**	-60 (-63)	122 **(120)**	-30 (-12)	1.55 **(1.56)**	1.50 (1.48)	1.45 (1.45)	1.36 (1.31)
II	16 or 17	**VII_r**	-60 (-65)	122 **(121)**	-30 (-12)	1.55 **(1.56)**	1.50 (1.48)	1.45 (1.44)	1.36 (1.31)
(b)B3LYP, neutral (cation)									
I	14	I_l	53 (75)	123 (135)	-170 (-150)	1.55 (1.59)	1.51 (1.49)	1.47 (1.44)	1.37 (1.32)
I	14	I_r	52 (69)	123 (135)	-170 (-154)	1.55 (1.59)	1.51 (1.48)	1.47 (1.44)	1.37 (1.32)
I	4 or 6	III_l	-64 (-84)	122 (136)	165 (-162)	1.55 (1.58)	1.51 (1.48)	1.47 (1.43)	1.37 (1.32)
I	4 or 6	III_r	-64 (-85)	122 (136)	165 (-162)	1.55 (1.58)	1.51 (1.48)	1.47 (1.43)	1.37 (1.32)
II	5 or 7	**II_l**	62 (62)	120 (131)	-4 **(-86)**	1.55 (1.57)	1.51 (1.49)	1.45 (1.44)	1.37 (1.32)
II	5 or 7	**II_r**	62 (63)	120 (130)	-4 **(-81)**	1.55 (1.57)	1.51 (1.49)	1.45 (1.44)	1.37 (1.32)
II	16 or 17	**VII_l**	-61 (-73)	122 **(140)**	-32 (-17)	1.57 **(1.64)**	1.51 (1.47)	1.45 (1.41)	1.37 (1.32)
II	16 or 17	**VII_r**	-61 (-72)	122 **(143)**	-32 (-18)	1.57 **(1.64)**	1.51 (1.47)	1.45 (1.41)	1.37 (1.33)

Table 4.4. Relative energies (in kcal/mol) of eight stable L-tyrosine conformers in the neutral (cationic) gound state

Sub.	Origin Bands	Conf.	Neutral		Cation	
			M05-2X	B3LYP	M052X	B3LYP
I	14	I_l	0.00 (1)	0.00 (1)	4.60	4.69
I	14	I_r	0.32 (2)	0.35 (4)	5.53	5.51
I	4 or 6	III_l	1.05 (6)	0.13 (3)	8.70	5.56
I	4 or 6	III_r	0.82 (5)	0.11 (2)	8.45	5.37
II	5 or 7	II_l	0.56 (3)	0.91 (8)	0.18	0.00
II	5 or 7	II_r	0.66 (4)	0.90 (7)	0.00	0.38
II	16 or 17	VII_l	1.86 (13)	1.15 (9)	0.98	0.18
II	16 or 17	VII_r	1.65 (11)	1.15 (10)	1.30	0.04

(3) L-tryptophan

For L-tryptophan, six conformers (A–F) have been experimentally identified.[20] Figure 4.3 shows the geometrical structures of L-tryptophan for the most six valid neutral conformers and five cationic ones in their ground states. The structures were calculated by using M05-2X optimization with the 6-311+G* basis set and they were tentatively classified into two subgroups, I (I, III) and II (II, V), as for L-phenylalanine.[16]

It has already been qualitatively understood that each subgroup for L-phenylalanine has two different types of intramolecular hydrogen bonding.[5]

The two functionals, M05-2X and B3LYP, gave almost the same results regarding geometrical structures for L-tryptophan in neutral and cation. The neutral conformers of six alphabetic capital letters, A, B, C, D, E, and F in Figure 4.3 were quoted from ref 20 in which their assignment was carried out using mass-selected resonant two photon ionization (MS-R2PI) spectra, UV–UV hole-burn spectra, and IR–UV ion dip spectra together with predictions by *ab* initio MO calculations employing B3LYP.

The four Roman numerals, I, II, III, and V, were adopted from ref 16 on L-phenylalanine for similarity of the backbone, that is, orientation of backbone (alanine) in L-tryptophan and L-phenylalanine are almost the same. Subscripts a and b represent different orientations of the residue (indole ring) because of asymmetry in L-tryptophan.

In Figure 4.3, M05-2X results show that cationic structures of conformer I_a belonging to subgroup I are similar to that of L-phenylalanine, in other words, geometry of the backbone (alanine) in L-tryptophan is almost same as that of L-phenylalanine, while conformer III_a in radical cations is quite different from that of L-phenylalanine.

Table 4.5 shows the calculated molecular constants for six conformers of L-tryptophan in the neutral and radical cationic ground states. Both the M05-2X and B3LYP results have the similar tendency in their molecular constants.

Furthermore, there exist no significant differences in the molecular constants between subgroups I and II. Cationic L-tryptophan conformers belonging to subgroup II have neither C_α–C_β elongation nor planarity of C_α–NH_2, which is different from cationic conformers belonging to subgroup II of L-phenylalanine.

Table 4.6 shows relative energies of L-tryptophan conformers in neutrals (cations), which were calculated by using M05-2X and conventional DFT (B3LYP) method including zero point energy corrections (ZPE_{corr}). The M05-2X functional gave more reasonable candidates corresponding to observed conformers in L-tryptophan than the B3LYP based on order of relative energies.

As mentioned above, there exist two types of hydrogen bonding in L-tryptophan based on their structures and frequency analysis and then conformer I_a and III_a, which belong to subgroup I and which have intramolecular hydrogen bonding (–COOH → $–NH_2$), should be candidates for observed conformers A and F.

Conformer II_a, II_b, V_a, and V_b, which belong to subgroup II and which have free –OH, should correspond to conformers E, C, B, and D, which were observed by Simons and co-workers in ref 20. The results from newly employed M05-2X for L-tryptophan in neutral provides the validity to show good agreement with assignments by Simon and co-workers.

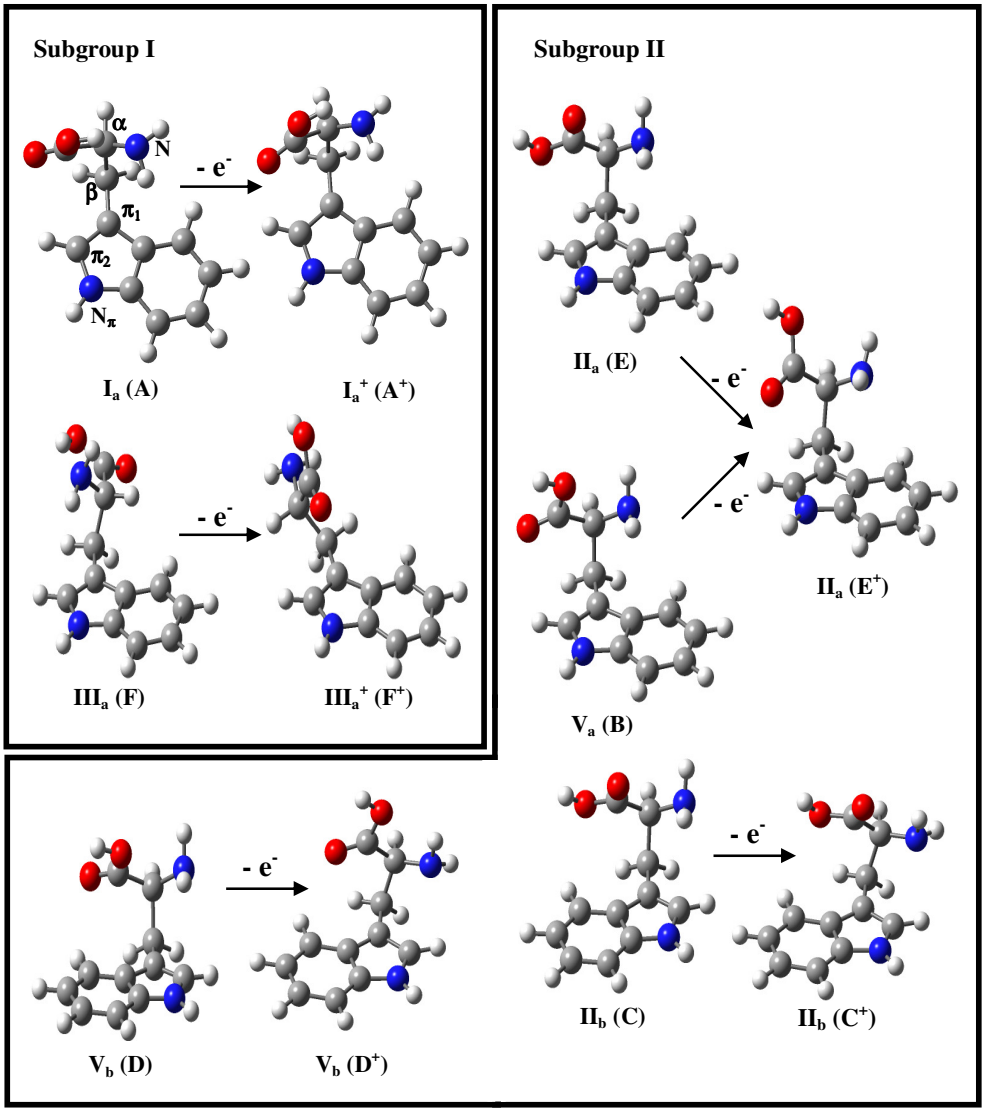


Figure 4.3. The six stable structures of representative L-tryptophan conformers in neutrals and cations.

Table 4.5. Molecular constants for six L-tryptophan conformers in neutrals and those for cations in parentheses.

Sub.	Experiment Assignment	Conf.	Dihedral angle (°)		Bond length (Å)			
			$C_{\pi1}$-C_β-C_α-N	C_α-NH_2	C_α-C_β	$C_{\pi1}$-$C_{\pi2}$	C_α-N	$C_{\pi2}$-N_π
(a) M05-2X, neutral (cation)								
I	A	I_a	55 (68)	122 (130)	1.53 (1.56)	1.37 (1.44)	1.46 (1.45)	1.37 (1.32)
I	F	III_a	-62 (172)	121 (128)	1.54 (1.54)	1.37 (1.43)	1.46 (1.45)	1.38 (1.32)
II	E	II_a	65 (58)	119 (123)	1.54 (1.54)	1.37 (1.43)	1.45 (1.45)	1.38 (1.32)
II	C	II_b	63 (63)	119 (122)	1.54 (1.55)	1.37 (1.43)	1.45 (1.45)	1.38 (1.32)
II	B	V_a	63 (58)	121 (123)	1.54 (1.54)	1.37 (1.43)	1.45 (1.45)	1.38 (1.32)
II	D	V_b	60 (59)	120 (120)	1.54 (1.54)	1.37 (1.43)	1.45 (1.46)	1.38 (1.32)
(b) B3LYP, neutral (cation)								
I	A	I_a	55 (72)	123 (131)	1.55 (1.57)	1.37 (1.43)	1.47 (1.45)	1.38 (1.33)
I	F	III_a	-62 (171)	121 (129)	1.55 (1.55)	1.37 (1.43)	1.47 (1.45)	1.38 (1.33)
II	E	II_a	65 (58)	120 (126)	1.55 (1.56)	1.37 (1.43)	1.45 (1.45)	1.38 (1.33)
II	C	II_b	64 (63)	119 (122)	1.55 (1.56)	1.37 (1.43)	1.45 (1.46)	1.38 (1.33)
II	B	V_a	64 (58)	121 (126)	1.55 (1.56)	1.37 (1.43)	1.46 (1.45)	1.38 (1.33)
II	D	V_b	60 (59)	121 (120)	1.55 (1.55)	1.37 (1.43)	1.45 (1.47)	1.38 (1.33)

Table 4.6. Relative energies (in kcal/mol) of six stable L-tryptophan conformers in the neutral (cationic) ground state.

Sub.	Experiment Assignments	Conf.	Neutral		Cation	
			M05-2X	B3LYP	M052X	B3LYP
I	A	I_a	0.00 (1)	0.00 (1)	4.14	4.12
I	F	III_a	1.39 (3)	0.13 (2)	4.36	2.55
II	E	II_a	1.32 (2)	0.91 (7)	0.00	0.00
II	C	II_b	1.66 (4)	0.90 (10)	0.26	0.98
II	B	V_a	1.83 (5)	1.15 (13)	0.00	0.00
II	D	V_b	2.11 (9)	1.15 (14)	0.23	0.94

4.3.2. Correspondence between the observed and computed conformers for L-phenylalanine

For L-phenylalanine, it has remained to be determined why conformers IV, V, and VIII, which were predicted by B3LYP calculations, have not been identified experimentally.

Conformer IV belonging to subgroup I has a relative energy (ΔE) of 0.17 kcal/mol (B3LYP) as shown in Table 4.2 and an activation energy E_a of 0.04 kcal/mol (see Figure 4.4(a)) in its neutral state, which is efficient for IV to be produced in a chamber under experimental conditions at temperatures of 170–190 °C (kT = 0.88 – 0.92 kcal/mol).

Conformer VIII belonging to subgroup II has also not been assigned experimentally, though B3LYP value of ΔE = 1.09 kcal/mol shown in Table 4.2 indicates the possibility of its observation. The M05-2X value of ΔE = 0.79 kcal/mol for conformer IV is very close to that of conformer III (B), ΔE = 0.62 kcal/mol. Here, entropy effects were neglected since the change in dihedral angle C_γ–C_α–C_β–C_π, which is associated with the backbone and residue rotation, is very small between the two conformers.

Figure 4.4a(b) show the results of intrinsic reacton coordinate (IRC) profiles for two conformerizations: III(B) – IV conformerization and VII(A) – VIII – IX(E) conformerization, respectively. Reaction paths through corresponding transition

state (TS) are determined by IRC. Here, zero-point energy correction (ZPE_{corr}) was taken into account. In Figure 4.4(a), III(B)–IV conformerization, both the M05-2X and B3LYP results gave almost the same behaviors for IRC profiles with a transition state (TS) having one imaginary frequency: the E_a of IV is 0.11 kcal/mol for M05-2X and 0.04 kcal/mol for B3LYP; ΔE between III(B) and IV is 0.17 kcal/mol for both M05-2X and B3LYP.

For III(B)–IV conformerization, reaction coordinate is rotation of C_α–C_α keeping intramolecular hydrogen bonding (–COOH → $–NH_2$). It has been shown that complex molecules with multiple degrees of conformational flexibility need a barrier height of 600 cm^{-1} (1.72 kcal/mol) or higher to be trapped in jet expansion experiments.[17]

This shows that under ordinary experimental conditions for a supersonic jet expansion experiment it is highly possible for III (B) and IV to be detected as the same conformer. The population ratio between III(B) and IV is roughly estimated to be 1:0.21 using a vibrational temperature of T = 57 K, which was adopted from rotational band-counter analyses of conformers in a supersonic jet expansion experiment.[5]

The ratio obtained by assuming that conformers retain thermal equilibrium is approximate, and detailed discussions are needed on cooling and trapping processes in supersonic jet expansion experiments in order to obtain more reliable population ratios.[17,18]

Although it is difficult to find clear evidence of the coexistence of the two conformers from the corresponding band in UV spectra, the possibility of coexistence is indicated by comparison of experimental rotational band contours with theoretically predicted results.[2] Here, the experimental band contours cannot be fitted well to the theoretical ones within one conformer III(B) model.

For 4.4(b) VII(A)–VIII–IX(E) conformerization, rotation of C_α–NH_2 is reaction coordinate. Two results are totally different: the B3LYP cannot explain the fact that VIII have not been observed. On the other hand, its M05-2X calculation showed ΔE = 2.06 kcal/mol ≈ 700 cm^{-1}, indicating that the energy is too high for it to be observed. The population of conformer VIII, which was estimated in the same way as described above, is too small to be observed, while VII(A) and IX(E) can be observed.

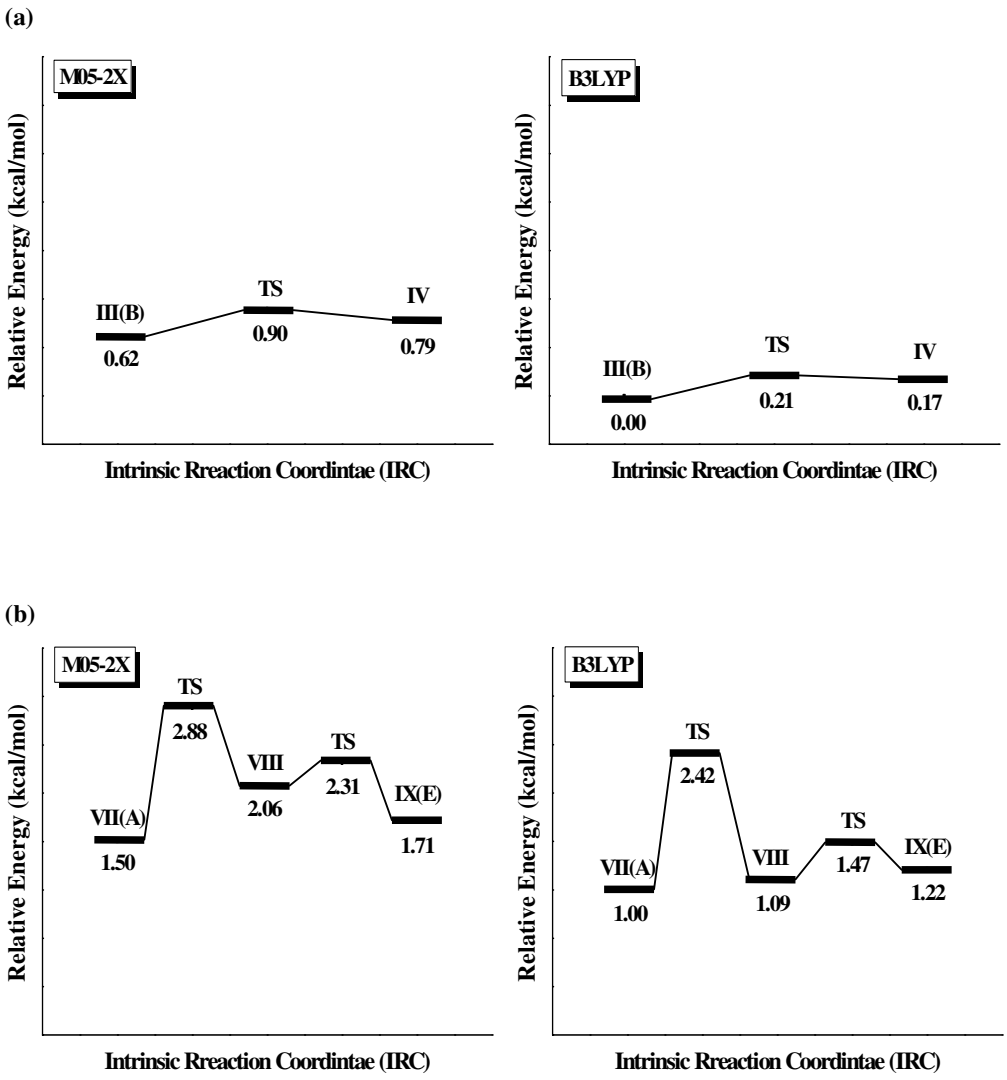


Figure 4.4. Intrinsic reaction coordinate (IRC) profiles on conformerizations of L-phenylalanine conformers in neutrals. (a) III(B)–IV conformerization from M05-2X and B3LYP; (b) VII(A)–VIII–IX(E) conformarization from M05-2X and B3LYP.

4.4. Conclusion

Firstly, B3LYP calculations were carried out in order to confirm the previous classification of the conformers into two subgroups (I and II) and then M05-2X calculations were conducted for both the neutral and the cationic conformers of all the aromatic amino acids. The two functionals, M05-2X and B3LYP, gave the same nine stable conformers in their neutral state.

However, quantitatively, their relative energies were different, the magnitudes of which are sufficient to identify the conformers: the M05-2X results can unambiguously explain the correspondence between the observed and the predicted conformers in terms of the relative energy and activation energy between the two conformers.

Especially, some cationic conformers were shown significantly different structure between two methods: for L-phenylalanine, conformer I and V in Figure 4.1a(b), for L-tyrosine, all conformers belonging to subgroup II in Figure 4.2(b). The difference results from absence of noncovalent intramolecular interaction in B3LYP.

By showing our M05-2X results, it is strongly emphasized to include long-range noncovalent effects for study of intramolecular hydrogen bonding systems like aromatic amino acids.

Here, it is newly suggested that Conformer V is of a hybrid character, belonging to subgroup I in its neutral form and belonging to subgroup II in its cation form. This character should be confirmed by experiments.

On the basis of the intrinsic reaction coordinate (IRC) results predicted by the new functional, it can be inferred why conformers IV, V, and VIII, which were predicted by B3LYP calculations, have not been identified experimentally.

4.5. References

[1] Zhao, Y; Truhlar, D. G. *J. Phys. Chem. A* 2004,*108,* 6908..

[2] Lee, Y.; Jung, J.; Kim, B.; Butz, P.; Snoek, L. C.; Kroemer, R. T.; Simons, J. P. *J. Phys. Chem. A* **2004**, *108,* 69.

[3] Lee, K. T.; Sung, J.; Lee, K. J.; Park, Y. D.; Kim, S. K. *Angew. Chem. Int. Ed.* **2002,** 41, 4114.

[4] Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321,* 49.

[5] Bouchoux, G.; Bourcier, S.; Blanc, V.; Desaphy, S. *J. Phy. Chem. B* **2009**, *113,* 5549.

[6] Lee, K. T.; Sung, J.; Lee, K. J.; Kim, S. K.; Park, Y. D. *Chem. Phys. Lett.* **2003**, 368, 263

[7] Lee, K. T.; Kim, H. M.; Han, K. Y.; Sung, J.; Lee, K. J.; Kim, S. K. *J. Am. Chem. Soc.* **2007**, *129*, 2588.

[8] Hashimoto, T.; Takasu, Y.; Yamada, Y.; Ebata, T. *Chem. Phys. Lett.* **2006**, *421,* 227.

[9] Huang, Z.; Yu, W.; Lin, Z. *Theochem-J. Mol. Struct.* **2006**, *758,* 195.

[10] Ebata, T.; Hashimoto, T.; Ito, T.; Inokuchi, Y.; Altunsu, F.; Brutschy, B.; Tarakeshwar, P. *Phys. Chem. Chem. Phys*. **2006**, *8,* 4783.

[11] Ebata, T.; Hashimoto, T.; Ito, T.; Inokuchi, Y.; Altunsu, F.; Brutschy, B.;

Tarakeshwar, P. *Phys. Chem. Chem. Phys*. **2006**, *8,* 4783.

[12] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[13] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[14] Gaussian 09, Revision A.1, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2009**.

[15] Huang, Z.; Yu, W.; Lin, Z. *Theochem-J. Mol. Struct.* **2006**, *758,* 195.

[16] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583.

[17] Godfrey, P. D.; Brown, R. D.; Rodgers, F. M. *J. Mol. Struct.* **1996**, 376, 65.

[18] Florio, G. M.; Christie, R. A.; Jordan, K. D.; Zwier, T. S. *J. Am. Chem. Soc.* **2002**, *125*, 10236.

[19] Inokuchi, Y.; Cobayashi, Y.; Ito, T.; Ebata, T. *J. Phys. Chem. A* **2007**, *111,* 3209.

[20] Snoek, L. C.; Kroemer, R. T.; Hockrige, M. R.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3,* 1819.

[21] Zhao, Y.; Truhlar, D. G. *Accounts Chem. Res.* **2008**, *41,* 157.

[22] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[23] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123,* 161103.

[24] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[25] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2011**, *115*, 9658.

Chapter 5.

A Tool Investigating

Conformation-Dependent Property:

Ionization Energy

5.1. Introduction

Ionization energy (IE) means the minimum quantity of energy required in order to remove one electron from an atom or molecule. This is the definition of adiabatic ionization energy. In the process of adiabatic ionization, molecular geometry should be changed.

Vertical ionization energy is defined as the energy gap between neutral and cation states under the nucleus frozen condition. It is concerned with Franck-Condon overlap and molecular geometry is fixed on the ionization. The definition of adiabatic and vertical ionization energies are illustrated in Figure 5.1.

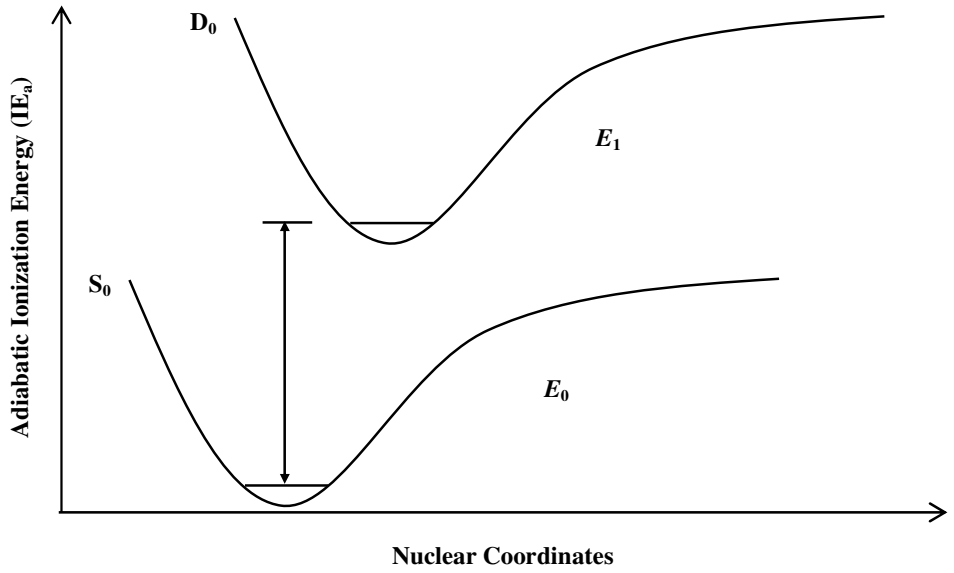
Photoionization, a kind of the ionization methods, is described to be easy understand in ref 1 and it utilizes electromagnetic radiation or photon in order to provide energy source to ionized neutral species.

Nowadays, a variety of spectroscopic techniques have been developed to ionize molecules and one of the most popular way is resonant two-photon ionization (R2PI) technique: the wavelength of the first photon is tuned to a resonant transition and the subsequent second photon brings the molecule above the ionization threshold.

By using this technique, the ion signal rise abruptly from zero, the onset of which can be regarded as the IE. This spectrum is called photoionization–efficiency (PIE) curve. The PIE curve of the result obtained for L-phenylalanine

shows that there is strong conformation-dependent correlation between the ionization energies of molecules and intramolecular hydrogen bonding interactions.[2]

In the next section, theoretical methods and procedures are briefly described. In section 5.3, vertical and adiabatic ionization energies of conformers for the three aromatic amino acids are calculated to provide further evidence to support the regrouping of subgroups I and II.



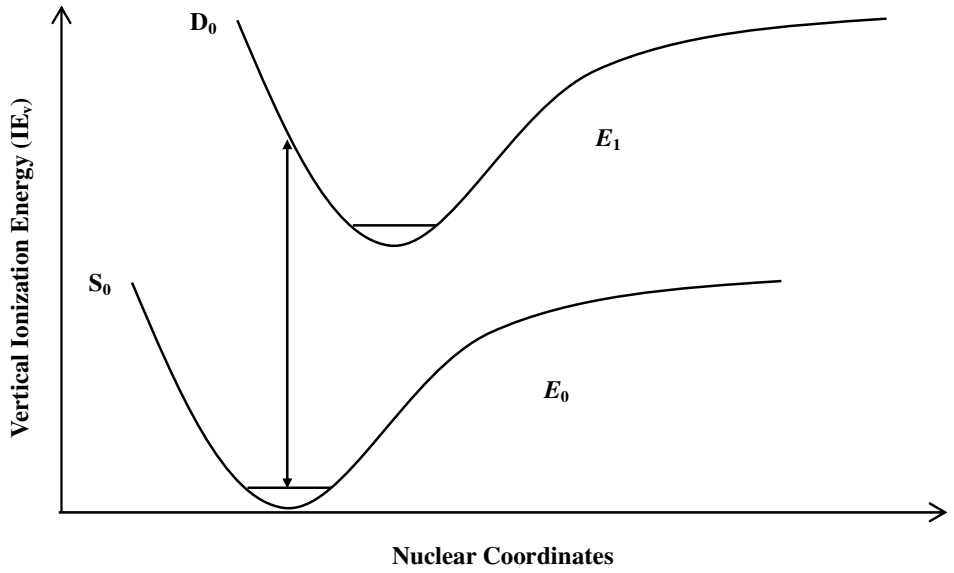


Figure 5.1. Ionization energy diagram.

5.2. Computational procedure

Firstly, geometry optimization is carried out for all neutral conformers of the three aromatic amino acids in ground state by setting the charge and the multiplicity to 0 and 1 at M05-2X/6-311+G* level of theory.

In the second place, it is evaluated by single point energy calculation for vertical ionization of each conformer at the optimized neutral in ground state by setting the charge and the spin multiplicity to +1 and 2 at M05-2X/6-311+G* level, respectively.

Whereas, geometry optimization for each cationic conformer of the three aromatic amino acids in their ground must be performed to evaluate the value of adiabatic ionization energy.

Here, all the calculations for ionization energy include zero-point energy corrections (ZPE_{corr}). The calculation at B3LYP/6-311+G* also is applied all calculation procedure for ionization energy with M05-2X method, simultaneously.

5.3. Results and discussion

5.3.1. L-phenylalanine

Figure 5.2a(b) shows calculated adiabatic (vertical) ionization energies of nine L-phenylalanine conformers. Both the adiabatic and vertical ionization energies calculated with M05-2X are higher than those calculated with B3LYP by ~0.5 eV for all of the conformers.

There have been no experimental reports on adiabatic ionization energies except for photoelectron measurements. This is because Franck-Condon factors for 0-0 transitions are negligibly small, which are due to large geometrical changes between the neutral and the cationic states, that is, elongation of the $C_\alpha-C_\beta$ bond and dihedral angle of $N-C_\alpha-C_\beta-C_\pi$ and pyramidal structure for $-NH_2$ on ionization as shown in Table 4.1.

Photoelectron spectroscopic values of toluene (8.82 eV)[5] and L-phenylalanine (8.50 eV)[6] are plotted in Figure 5.2(a). The predicted adiabatic ionization energies, I_a (M05-2X and B3LYP) for toluene are also plotted to estimate a possible error in the calculations. The M05-2X adiabatic ionization energies of the subgroup I conformers, I (X), III(B), and IV, are close to that of toluene, and the adiabatic ionization energies of the subgroup II conformers, VII(A), VI(C), II(D), and IX(E), are close to that of L-phenylalanine by photoelectron measurement, and

their energy differences are within 0.1eV. The adiabatic energies calculated B3LYP, on the other hand, are far from the experimental values by ~0.6 eV for the subgroup I conformers and ~0.4 eV for the subgroup II conformers.

Concerning vertical ionization energies, conformational-dependent ionization energies have been reported by *Lee* **et al.**[7] They measured photoionization efficiency (PIE) curves as a function of the wavelength of the ionization laser, while fixing the wavelength of the excitation laser at the S_1 band origin of each conformer in mass-resolved resonant two-photon ionization (R2PI) experiments.

The ion signal in PIE curve appears at ~8.8 eV for all the subgroup II conformers (VII(A), VI(C), II(D), and IX(E)) and at ~9.0 eV for I(X) and 9.15 eV for III(B) of subgroup I conformers. These correspond to the vertical ionization energies. The values thus estimated are indicated by the solid squares in Figure 5.2(b). Only one experimental value of 9.4 eV, which was obtained from a photoelectron spectrum[8] of a mixture of L-phenylalanine conformers, is given.

The estimated values for the vertical ionization energies of subgroup I conformers are marked by solid squares with an arrow, which represent lower bounds for the vertical ionization energies observed because of restriction in the wavelength ranges in the PIE measurement.

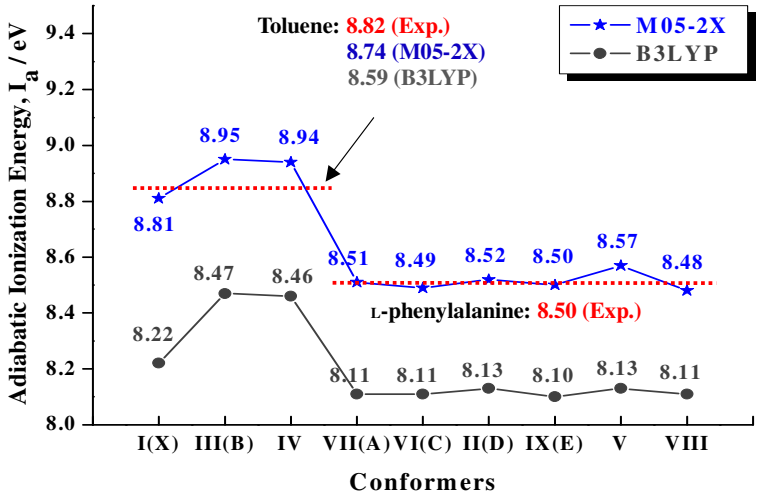
It is demonstrated that the M05-2X functional satisfactorily reproduces observed vertical ionization energies, being close to both the PIE and the

photoelectron experimental values, while the B3LYP functional could not quantitatively reproduce these experimental values.

It has already been pointed out that there is a tendency in ionization energy differences between subgroup I and II conformers, that is, the ionization energies of subgroup I conformers are higher than those of subgroup II.

It is important that ionization energy can be utilized as a tool to verify the conformation-dependent properties. As shown in Figure 5.2a(b) (M05-2X) and ref 7, one can understand there exists different type of structure between conformers of L-phenylalanine based on their tendencies of energy gap.

(a)



(b)

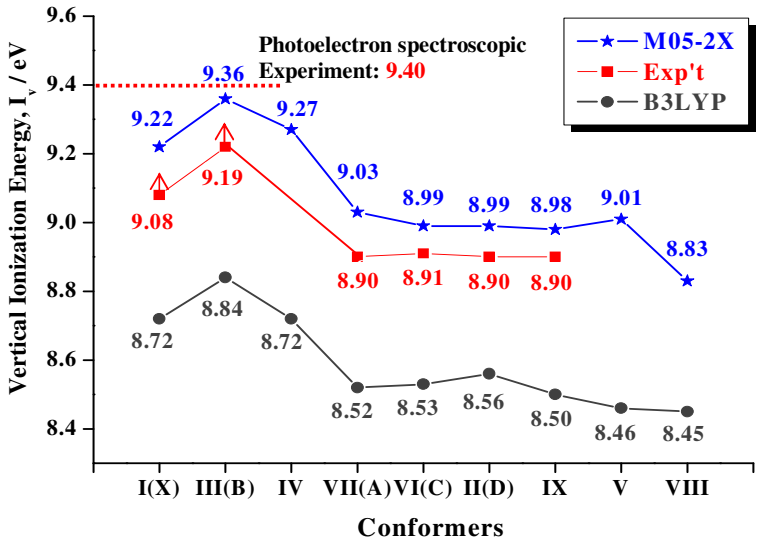


Figure 5.2. M05-2X and B3LYP results for (a) adiabatic and (b) vertical ionization energies for the nine most stable conformers of L-phenylalanine.

5.3.2. L-tyrosine and L-tryptophan

Figure 5.3 shows the calculated adiabatic ionization energies for L-tyrosine in the middle and those for L-tryptophan at the bottom. For comparison, in the upper part, adiabatic ionization energies for L-phenylalanine[3] are also shown. The adiabatic ionization energies calculated by M05-2X are higher than those calculated by B3LYP for all of the conformers by nearly 0.4 eV. Here, it can be seen that adiabatic ionization energies for all of the conformers of both L-tyrosine and L-tryptophan can be classified into two groups depending on subgroups I and II as those for L-phenylalanine: adiabatic ionization energies for the subgroup I conformers are higher than those for the subgroup II conformers. The energy difference between subgroup I and subgroup II becomes small, i.e., 0.29, 0.21, and 0.19 eV, as aromatic amino acid is replaced from L-phenylalanine to L-tyrosine and L-tryptophan.

Experimental values of adiabatic ionization energy are also shown in Figure 5.3. The adiabatic ionization energies of subgroup II (II and VII) conformers of L-tyrosine, which were calculated by M05-2X, reproduces those observed by photoelectron spectroscopy on L-tyrosine,[6] whereas for conformer I and III belonging to subgroup I, their adiabatic ionization energies reproduce photoelectron spectroscopic values of cresol (8.35 eV)[5] and L-tyrosine (8.00 eV)[6] are plotted in the middle of Figure 5.3.

Both the M05-2X and the B3LYP values of adiabatic ionization energy for cresol are also plotted to estimate a possible error in the calculations.

For L-tryptophan, the M05-2X adiabatic ionization energies of the subgroup I conformers, I_a (A) and III_a (F), are close to that of skatol at the bottom of Figure 5.3. The adiabatic ionization energies of the subgroup II conformers, II_a (E), II_b (C), V_a (B), and V_b (D), are close to that of L-tryptophan and their energy differences are within 0.10 eV.[10]

Alanine is the backbone common to all of the three aromatic amino acids, and the experimental values of their adiabatic ionization energies are 8.82, 8.88, 8.35, and 7.51 eV for toluene,[5] alanine,[6] cresol,[11] and skatol,[12] respectively.

The ionization energies predicted by M05-2X (B3LYP) with the 6-311+G* basis set are 8.74 (8.59), 8.12 (7.95), and 7.53 (7.31) eV for toluene, cresol, and skatol, and 9.36 (9.07), 8.94 (8.70) eV for alanine belonging to subgroups I and II, respectively. Photoelectron spectroscopic values of skatol (7.51 eV)[12] and L-tryptophan (7.30 eV)[10] are plotted in the bottom of Figure 5.3. Both the M05-2X and B3LYP values of adiabatic ionization energy for skatol are also plotted to estimate a possible error in their calculations.

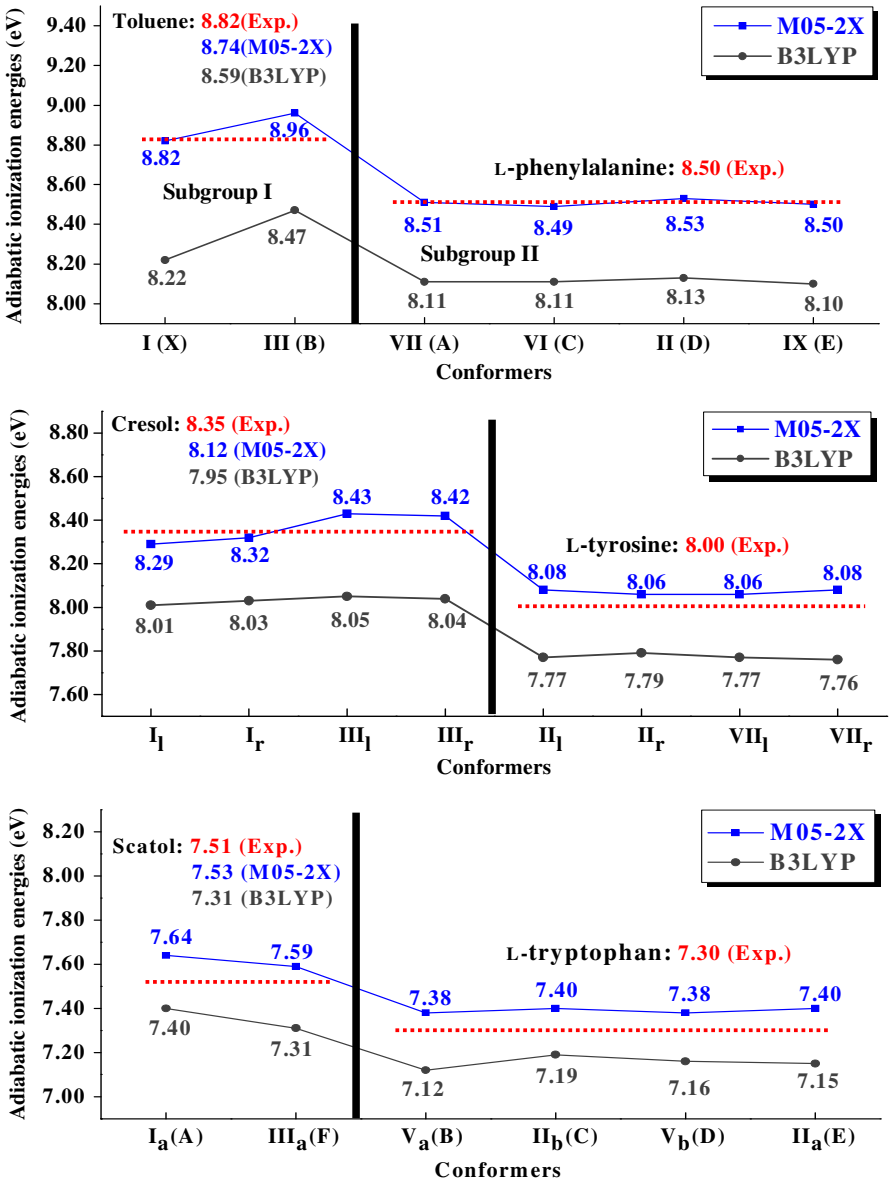


Figure 5.3. M05-2X and B3LYP results for adiabatic ionization energies of the six stable conformers in L-phenylalanine (upper), eight stable conformers in L-tyrosine (middle), and six stable conformers in L-tryptophan (lower).

Figure 5.4 shows a qualitative explanation for the calculated adiabatic ionization energies of each aromatic amino acid based on a simple molecular orbital picture.

Figure 5.4(a) shows the ionization of the three aromatic amino acids belonging to subgroup I. The tendency can be explained in terms of intramolecular hydrogen bonding between lone pair electrons of the amino group and hydrogen of the carboxyl group for subgroup I conformers. The n-MO of the amino group is stabilized as a result of hydrogen bonding, and interactions between the π-HOMOs of the phenyl group and the n-MO become negligibly small.

Figure 5.4(b) shows the ionization of the aromatic amino acids belonging to subgroup II. In this case, on the other hand, ionization takes place from the HOMO of aromatic amino acids.

In this section, I try to discuss the characteristic only shown in subgroup II of L-phenylalanine using a model in 5.4(b). Only subgroup II of L-phenylalanine, the energy levels of π-MOs and n-MO are close to each other as can be seen in Figure 5.4(b), and a pair of MOs is formed between a π-MO and an n-MO.

It results from their submoiety having almost same ionization energy: (toluene: 8.82 eV, alanine: 8.88eV) as mentioned above. This newly formed MOs are called hyperconjugative MOs.[9] Ionization from one of the hyperconjugative MOs with an antibonding character results in lowering of ionization energies of subgroup II conformers. C_α–C_β elongation of the subgroup II conformers on ionization can

qualitatively be explained in terms of electron dynamics associated with hyperconjugative MO.

The C_α–C_β elongation is mainly brought about by an electron release from the antibonding hyperconjugative MO: after ionization, both chromophores, phenyl and amino groups, have a positive partial charge with equal weights and attract electrons in the C_α–C_β bond. As a result, elongation occurs.

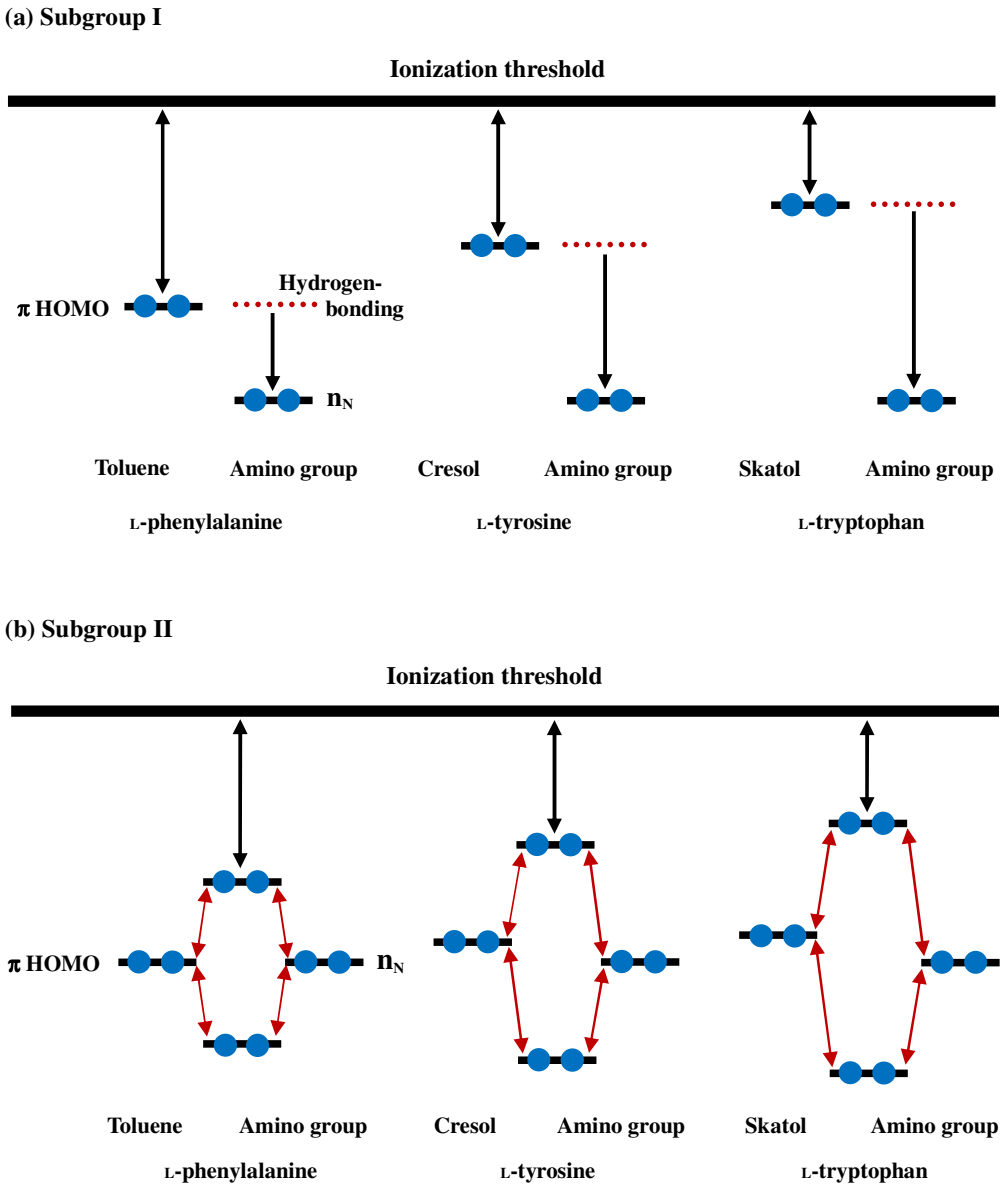


Figure 5.4. An MO picture showing the differences in ionization energies between conformers belonging to subgroup I and those belonging to subgroup II for the three aromatic amino acids

5.4. Conclusion

In this chapter, M05-2X calculations gave adiabatic (vertical) ionization energies in cationic conformers for all the three aromatic amino acids and several chromophores in order to explain the correlation between the conformational dependent properties of molecule and ionization energies, theoritically.

Based on the M05-2X results, unique character of subgroup II for L-phenylalanine could be describe by employing a simple MO model. Figure 5.4 can explain the reason why only L-phenylalanine conformers belonging to subgroup II against other aromatic amino acids (L-tyrosine and L-tryptophan) shows significant elongated bond length of the C_α–C_β and planarity of $-NH_2$ when they became cations.

The present results provide a reliable basis for analysis of ionization efficiency spectra and investigation of mechanisms of chemical reactions of the aromatic amino acids. Ionization-induced conformational change is one of the examples of chemical reactions.[13,14]

5.5. References

[1] Lee, K. T. Ph. D. Dissertation, Seoul National University, Seoul, 2002. pp.20-24.

[2] Lee, K. T.; Sung, J.; Lee, K. J.; Park, Y. D.; Kim, S. K. *Angew. Chem. Int. Ed.* **2002**, *41,* 4114.

[3] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583

[4] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2011**, *115*, 9658.

[5] Lu, Kuelh-Tzu; Eiden, G. C.; Weisshaar, J. C. *J. Phys. Chem.* **1992,** *96,* 9742.

[6] Campbell, S.; Beauchamp, J. L.; Rempe, M.; Lichtenberger, D. L. *Int. J. Mass Spectrom. Ion Processes* **1992,** *117*, 83.

[7] Lee, K. T.; Sung, J.; Lee, K. J.; Park, Y. D.; Kim, S. K. *Angew. Chem. Int. Ed.* **2002**, *41,* 4114.

[8] Cannington, P. H.; Ham, N. S. *J.Electron Spectrosc. Relat. Phenom.* **1983**, 32, 139.

[9] Sahnoun, R.; Fujimura, Y.; Kabuto, K.; Takeuchi, Y.; Noyori, R. *J. Org. Chem.* **2007**, *72,* 7923.

[10] Wilson, K. R.; Jimenez-Cruz, M., Nicolas, C.; Belau, L.; Leone, S. R.; Ahmed, M. *J. Phys. Chem. A* **2006**, *110,* 2106.

[11] Palmer, M. H.; Moyes, W.; Speirs, M.; Ridyard, J. N. A. *J. Mol. Struct.* **1979**, *52*, 293.

[12] Hager, J. W.; Wallace, S. C. *Anal. Chem.* **1988**, *60*, 5.

[13] Miyazaki, M.; Fujii, A.; Ebata, T.; Mikami, N. *Phys. Chem. CHem. Phys.* **2003**, *5*, 1137.

[14] Ishiuchi, S. –i.; Sakai, M.; Tsuchida, Y.; Takeda, A.; Kawashima, Y.; Fujii, M.; Dopfer, O.; Müller-Dethlefs, K. *Angew. Chem. Int. Ed.* **2005**, 44, 6149.

Chapter 6.

Hyperconjugation Effects on Charge Distribution in the Cationic Aromatic Amino Acids

6.1. Introduction

Weak interaction dominated by dispersion and intramoleular hydrogen bonding are important for van der Waals molecules and biological systems. The molecular shape and conformation are controlled by covalent bonds and nonbonded interactions, operating through space between neighboring groups of atoms or local charges within the molecules.[1]

Weinkauf **et al.** suggested that charge localization in the peptide be achieved by resonant UV two-photon ionization at an aromatic chromophore.[2] In other words, resonant absorption of a first visible photon activates charge flow in the cations and that of a second one probes presence of charge in the aromatic chromophore.

According to ref 3, a surplus charge in flexible molecules can influence the conformational landscape. For the reason, there exists correlation among structure, charge and reactivity of biomolecules, namely amino acids and proteins.

A new grouping of the stable conformers into two subgroups I and II is given based on the calculated results together with the results of the charge distribution in cations described in section 6.3. Delocalization of charge distributions between the amino and phenyl groups in L-phenylalanine on ionization is explained in terms of "hyperconjugation" in which the highest occupied molecular orbital (HOMO) of the phenyl group and the nonbonding molecular orbital (n-MO) of the

amino group form new bonding and antibonding MOs.

The charge distribution for cations is obtained using natural bond orbital[4] (NBO) analysis. The charge distribution between the subgroups of L-tyrosine and L-tryptophan are investigated in addition to L-phenylalanine.

Accurate evaluation of conformation-dependent charge distributions is very important in studying the mechanisms of ionization-induced chemical reactions as well as determination of conformers.[3,5-7]

This chapter explains the different trend of charge distribution among the three aromatic amino acids by using the ab initio MO method[8] and DFT method at the M05-2X level of theory[9-11] taking into account the effect of noncovalent, long-range potential.

6.2. Computational procedure

All the calculations are carried out including zero-point energy corrections (ZPE_{corr}). Each conformer for all the three amino acids is optimized in the corresponding neutral ground state by setting the charge and the spin multiplicity to 0 and 1, respectively at M05-2X/6-311+G* level of theory. Their optimized geometries in neutral ground state are used as initial geometry for each cationic conformer based on Franck-Condon principle and then they are reoptimized by setting the charge and the spin multiplicity to 1 and 2, respectively, at M05-2X with 6-311+G* basis set in the same way as geometry optimization for the neutral conformers.

For investigation of their charge distribution, natural bond orbital (NBO) analysis is carried out. The charge analysis can be obtained by subtraction between the partial charges of cation and neutral, and the value of charge distribution from M05-2X result can be shown using GaussView5 without difficulty. The checkpoint file (*.chk), which is created as normal termination, should be converted to its formatted checkpoint file (*.fchk) in order to visualize the shape of molecular orbital for the optimized structure in the ground state.[4] B3LYP calculation also were accomplished with 6-311+G* basis set for comparison with the M05-2X results.

6.3. Results and discussion

In Figure 6.1, toluene, cresol, and skatol are employed as representative aromatic residue chromophores for L-phenylalanine, L-tyrosine, and L-tryptophan, respectively.

A typical molecule for charge doubly localization (division) upon ionization is 2-phenylethylamine[3] (2-PEA), analogue of phenylalanine, which is supposed to be formed by ethylbenzene as a residue and ethylamine as a backbone. (See Figure 6.1) According to ref 3, they have almost the same ionization energy (8.77 eV for ethylbenzene, 8.80 eV for ethylamine) in experiment .

Weinkauf **et al**. suggested that the charge be founded to be delocalized equally over the two sites in the case of similar ionization energies and the charge delocalization effect can be observed through “bond” or “space”.

For the chromophores of those aromatic amino acids, we reported that only toluene and alanine belonging to subgroup II have the similar adiabatic ionization energy.[6,7] In brief, the adiabatic ionization energies by M05-2X are 9.36, (8.94), 8.74, 8.12, and 7.53 eV for alanine belonging to subgroup I (II), toluene, cresol, and skatol, respectively.

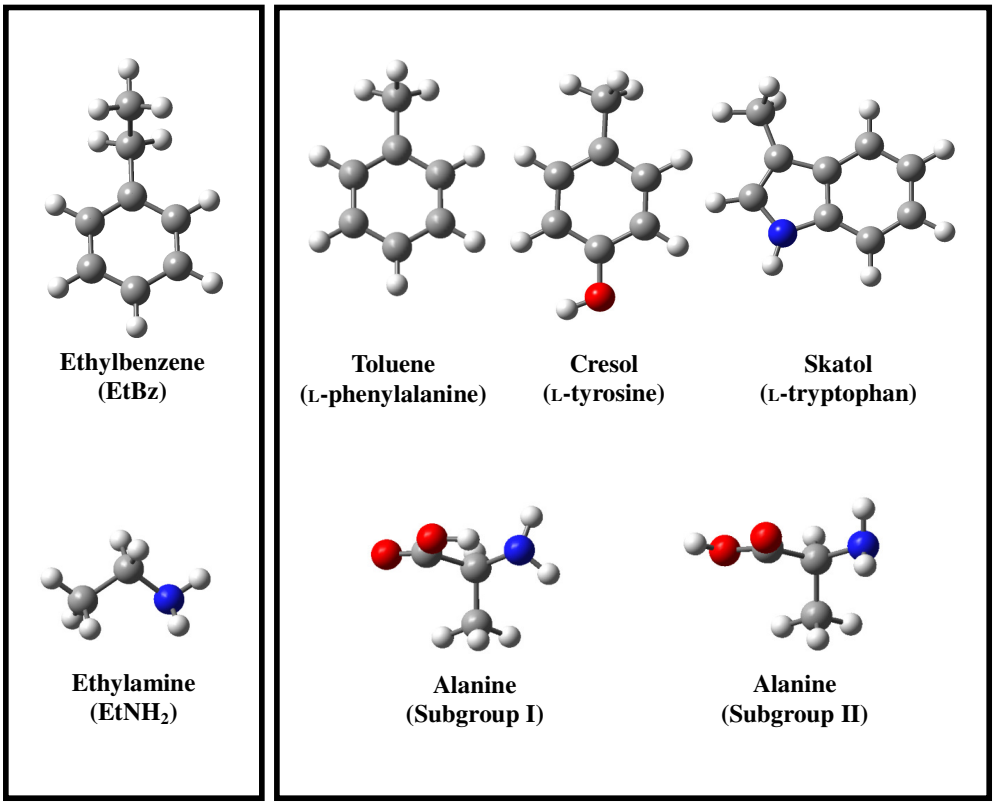


Figure 6.1. Several chromophores for explanation of charge distribution of the three aromatic amino acids.

6.3.1. L-phenylalanine

As mentioned above, remind the correlation between structure and charge of amino acids. Conformers of L-phenylalanine belonging to a given subgroup share the same characteristic features in the cationic ground state, that is, their structure for subgroup II have amino group's planarity, the elongated C_α–C_β bond length, and the cationic charge delocalized at two sites.

Whereas, those of subgroup I have the pyramidal amino group, normal C_α–C_β bond length, and the cationic charge localized in one region. A theoretical support for these findings has been recently given on the basis of M05-2X calculations.[6]

Table 6.1 shows the M05-2X results for the charge distributions on two chromophores, the phenyl ring and amino groups of the nine L-phenylalanine conformers. For comparison, B3LYP results also are listed in parentheses. Both the M05-2X and B3LYP results indicate that the charge distributions of conformers belonging to subgroup I are mostly localized on the phenyl ring.

However, for subgroup II conformers except conformer V, they are qualitatively delocalized, that is, the charges are shared between the two chromophores. The newly calculated value of 0.81 for conformer V strongly suggests that the cationic conformer belongs to subgroup I. In addition, the newly obtained results for the charge distributions show remarkable contrast in charge

distributions between the two subgroups compared with those obtained by the B3LYP method. In chapter 4, conformer V on ionization has already been verified as a hybrid type conformer based on structural change from M05-2X result.

In particular, charges of subgroup II conformers are almost equally distributed between the phenyl and the amine groups. The differences in charge distributions between the two subgroups were explained by the form of their intramolecular hydrogen bonding:[12] $-COOH \rightarrow -NH_2$ vs. $-NH_2 \rightarrow -OCOH$.

In addition to intramolecular hydrogen bonding ($-COOH \rightarrow -NH_2$) for conformers [I(X) and III(B)] belonging to subgroup I, we should stress that hyperconjugation in conformers [VII(A), VI(C), II(D), and IX(E)] belonging to subgroup II is responsible for the delocalized charges as qualitatively indicated in Table 6.1.

Figure 6.2 shows the phases of HOMO and HOMO-1 of the observed conformers: It can be seen that for subgroup II hyperconjugative HOMOs are formed through space interactions between the n-MO of the amino group and the π orbital of the phenyl ring [VII(A), IX(E)] or through bond interactions [(VI(C), II(D)].

Table 6.1. Partial charge distributions of cationic L-phenylalanine conformers.

Subgroup	Experimental Assignment	Conformer	M05-2X (B3LYP)			
			Phenyl group		Amino group	
I	X	I	0.79	(0.62)	0.04	(0.09)
I	B	III	0.73	(0.53)	0.09	(0.17)
I		IV	0.73	(0.54)	0.09	(0.16)
II (I)		**V**	**0.81**	**(0.42)**	**0.06**	**(0.33)**
II	A	VII	0.34	(0.41)	0.34	(0.35)
II	C	VI	0.38	(0.44)	0.33	(0.34)
II	D	II	0.39	(0.44)	0.31	(0.34)
II	E	IX	0.34	(0.41)	0.34	(0.35)
II		VIII	0.33	(0.40)	0.34	(0.35)

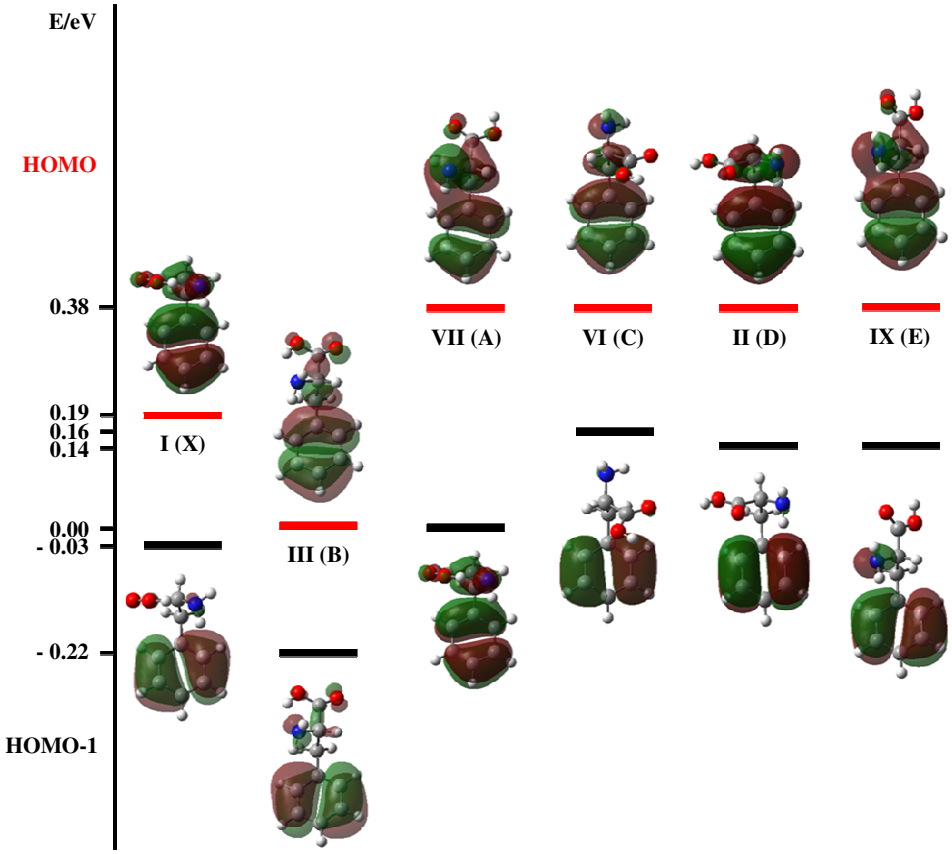


Figure 6.2. Molecular orbital phases and energies for HOMOs and HOMO-1s of six L-phenylalanine conformers.

6.3.2. L-tyrosine and L-tryptophan

Results of charge distribution obtained by natural bond orbital (NBO) analysis are given in Table 6.2. a(b) for L-tyrosine and L-tryptophan, respectively. It can be seen from the M05-2X results that their charge of 90% is localized on the residue (phenol and indole) for all of the L-tyrosine and L-tryptophan conformers belonging to subgroup I. The B3LYP results are marked in parentheses.

Note their charge localization for all the conformers belonging to subgroup II of L-tyrosine and L-tryptophan.[7] The values of their partial charge distribution are definitely equal to those of conformers belonging to subgroup I.

This makes clear contrast with L-phenylalanine conformers belonging to subgroup II, in which charge is doubly localized (divided) in each phenyl group and amino group with equal weight.[6] It is considered that the different tendency for charge localization between L-phenylalanine vs. L-tyrosine (L-tryptophan) is concerned with their adiabatic ionization energies predicted in the previous chapter.

Table 6.2. Partial charge distributions of L-tyrosine and L-tryptophan conformers in cations.

(a) L-tyrosine

Subgroup	Experiment Assignments	Conformer	M05-2X (B3LYP)			
			Phenol group		**Amino group**	
I	14	I_l	0.89	(0.78)	0.03	(0.07)
I	15	I_r	0.89	(0.78)	0.03	(0.08)
I	6	III_l	0.85	(0.69)	0.04	(0.10)
I	4	III_r	0.85	(0.69)	0.04	(0.10)
II	16	II_l	0.90	(0.75)	0.02	(0.10)
II	17	II_r	0.90	(0.75)	0.02	(0.09)
II	7	VII_l	0.91	(0.67)	0.01	(0.17)
II	5	VII_r	0.91	(0.65)	0.01	(0.18)

(b) L-tryptophan

Subgroup	Experiment Assignments	Conformers	M05-2X (B3LYP)			
			Indole group		**Amino group**	
I	A	I_a	0.90	(0.85)	0.02	(0.04)
I	F	III_a	0.87	(0.81)	0.04	(0.05)
II	E	II_a	0.89	(0.84)	0.03	(0.06)
II	C	II_b	0.92	(0.90)	0.01	(0.01)
II	B	V_a	0.90	(0.84)	0.03	(0.06)
II	D	V_b	0.93	(0.92)	0.01	(0.01)

Let us now make a semiquantitative discussion about the charge distribution shown in Table 6.2a(b) by using a simple two-state model introduced in Figure 5.4. According to the model, it consists of the HOMO of aromatic residue, ϕ_π, and nitrogen nonbonding MO of the backbone, ϕ_n. The HOMO (MO from which electron is released) of the aromatic amino acid $\Phi_n$can be expressed by a linear combination of these two MO's as $\Phi = \beta\,\phi_\pi - \alpha\phi_{nN}$. α and β are coefficients given as

$$\alpha = \frac{\left\{(\Delta\varepsilon^2 + 4\gamma^2)^{\frac{1}{2}} - \Delta\varepsilon\right\}}{\left[\left\{(\Delta\varepsilon^2 + 4\gamma^2)^{\frac{1}{2}} - \Delta\varepsilon\right\}^2 + 4\gamma^2\right]^{\frac{1}{2}}}$$

and

$$\beta = \frac{2|\gamma|}{\left[\left\{(\Delta\varepsilon^2 + 4\gamma^2)^{\frac{1}{2}} - \Delta\varepsilon\right\}^2 + 4\gamma^2\right]^{\frac{1}{2}}} \quad .$$

Here, $\Delta\varepsilon = \varepsilon_\pi - \varepsilon_{nN}$ and $\gamma = \langle\, \phi_\pi \big| \hat{H} \big| \phi_{nN} \rangle$; γ is the interaction energy between ϕ_π and ϕ_{nN}. $\hat{H}$ is the electronic Hamiltonian of aromatic amino acid at the ground state geometrical structure.

For L-phenylalanine belonging to subgroup II, $|\alpha| = |\beta|$ since both $\phi_\pi$and ϕ_{nN} have the same value of orbital energy and $\Delta\varepsilon = 0$. This means that charge is delocalized (divided) between the backbone and aromatic residue with equal weight, i.e., $|\alpha|^2 = |\beta|^2$.

In this case, γ can be estimated as the ionization energy difference between phenyl (or amino) group and L-phenylalanine. γ is estimated to be 0.32 eV from the experimental ionization energies.

In sequence, let us consider charge distribution of L-tyrosine by using the two-state model in Figure 5.4. Coefficients α and β were estimated to be $|\alpha|^2 = 0.11$ and $|\beta|^2 = 0.89$. Here, we adopt $\gamma = 0.32$ eV, which is estimated from L-phenylalanine since the aromatic ring of L-tyrosine has the similar aromatic ring as that of L-phenylalanine. It should be noted from comparison of geometrical structures between L-tyrosine and L-phenylalanine that their conformers have one-to-one correspondence with respect to their conformation.[7]

Furthermore, the angle between the directions of two orbitals ($\phi_\pi$and ϕ_{nN}) of each L-tyrosine conformer is almost equal to that of the corresponding L-phenylalanine conformer. Here, $\Delta\varepsilon$ is calculated as 0.82 eV and it is taken from the theoretical ionization energies of two substructures, cresol and alanine belonging to subgroup II. The values of the charge distribution obtained by using the simple model semiquantitatively explain those shown in Table 6.2(a).

In the similar way, we obtain $|\alpha|^2 = 0.44$ and $|\beta|^2 = 0.96$ for L-tryptophan with $\Delta\varepsilon = 1.41$ eV. It can be seen that there is significant difference in the charge distribution between the *ab* initio calculation and the simple two-state model. The main difference originates from use of the same interaction energy $\gamma = 0.32$ eV as that used in L-phenylalanine for evaluation of L-tryptophan charge distribution

in the simple model even though there is a large discrepancy in aromatic rings between L-tryptophan and L-phenylalanine. Actually, $\gamma = 0.53\text{eV}$ is obtained from the expression for $\alpha(\beta)$ with *ab* initio values of the charge distribution for L-tryptophan in Table 6.2(b).

On the other hand, $\gamma = 0.31$ eV is obtained from the expression for $\alpha(\beta)$ with *ab* initio values of the charge distribution for L-tyrosine in Table 6.2(a) and there is almost no discrepancy of γ between L-phenylalanine and L-tyrosine.

In summary, charge delocalization for the subgroup II conformers of L-phenylalanine can be considered as a special case. The tendency of no C_α–C_β elongation and nonplanarity of the C_α–NH_2 for all of the L-tyrosine and L-tryptophan conformers in radical cations, as shown in Tables 4.3 and 4.5, can be qualitatively explained using the same model.

Charge localization of aromatic residue in aromatic amino acids means that ionization takes place from the HOMO of aromatic residue. Therefore, the degree of contribution of molecular orbital character like nN and π plays an important role in determining whether the charge is localized or delocalized.

6.4. Conclusion

Conformation-dependent properties of the aromatic amino acids L-tyrosine and L-tryptophan against L-phenylalanine in radical cations have been studied by using new density functional theory, M05-2X. The results showed that all of the conformers of both L-tyrosine and L-tryptophan undergo charge localization respective of formation of intramolecular hydrogen bonding.

Such localization behavior originates from the existence of HOMO energy difference between the backbone and residue of the aromatic amino acid. Among the three aromatic amino acids, L-phenylalanine, L-tyrosine, and L-tryptophan, only subgroup II conformers of L-phenylalanine without intramolecular hydrogen bonding (–COOH → $–NH_2$) in the backbone undergo charge delocalization.

This is due to nearly the same energy values of HOMO's between the backbone, alanine, and aromatic residue, toluene. The analysis given here provides clues to analyze the dynamics of charges in peptides and amino acids.

6.5. References

[1] Robertson, E. G.' and Simons, J. P. *Phys. Chem. Chem. Phys.*, **2001**, *3*, 1

[2] Weinkauf, R.; Schanen, P.; Metsala, A.; Schlag, E. W.; Burgle, M.; Kessler, H. *J. Phys. Chem.* **1996**, *100*, 18567.

[3] Weinkauf, R.; Lehrer, F.; Schlag, E. W.; Metsala, A. *Faraday Discuss.* **2000**, *115*, 363.

[4] Gaussian 09, Revision A.1, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc.,

Wallingford CT, **2009**.

[5] Lee, K. T.; Kim, H. M.; Han, K. Y.; Sung, J.; Lee, K. J.; Kim, S. K. *J. Am. Chem. Soc.* **2007**, *129*, 2588.

[6] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2010**, *114*, 7583.

[7] Baek, K. Y.; Hayashi, M.; Fujimura, Y.; Lin, S. H.; Kim, S. K. *J. Phys. Chem. A* **2011**, *115*, 9658.

[8] Sahnoun, R.; Fujimura, Y.; Kabuto, K.; Takeuchi, Y.; Noyori, R. *J. Org. Chem.* **2007**, *72,* 7923.

[9] Zhao, Y; Truhlar, D. G. *J. Phys. Chem. A* **2004,***108,* 6908.

[10] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364.

[11] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120,* 215.

[12] Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321,* 49.

국문 초록

원거리 비공유결합 상호작용인 분자 내 수소결합을 하는 분자 시스템에 최적화된 새로운 밀도함수이론 (density functional theory) 방법인 M05-2X 을 이용하여 중성 및 양이온 상태에서의 세 가지 방향족 아미노산 (페닐알라닌, 타이로신, 트립토판)의 구조-의존적 특성을 연구하였다.

우선, M05-2X 방법의 신뢰도를 입증하기 위해서 몇 가지 발색단에 대하여 이 방법을 적용하였다. 또한 M05-2X 계산 결과와 비교하기 위해서 보편적으로 사용되는 밀도함수이론 방법인 B3LYP 계산도 병행하였다. M05-2X 로부터 예측된 이론값과 실험적으로 관측된 값은 오차범위 내에서 일치하였으나 B3LYP 의 계산결과는 실험치와 상당한 차이가 있음을 보였다. M05-2X 은 이들 발색단 중에서도 특히 방향족 아미노산에 대한 이온화 에너지 계산에서 실험치와 정확히 일치함을 보였는데, 이 결과는 M05-2X 가 분자 내 수소결합을 하는 생체 분자의 이론적 연구에 있어 가장 적합한 방법이라는 것을 증명한 것이라 하겠다.

본 연구에서의 모든 계산은 두 가지 계산방법 (M05-2X/6-311+G*와 B3LYP/6-311+G*)을 모두 적용하였으며, 영점 에너지 보정을 하였다. 중성과 양이온의 바닥상태에서 세 가지 방향족 아미노산의 구조 이성질체들에 대한 구조 최적화와 이온화 에너지, 전하 분포 계산을 수행하였다. 이론적으로는 예측 가능하지만 실험적으로 관측되지 않는 페닐알라닌의 몇몇 구조 이성질체에 대해 구조 이성질화 과정을 고유 반응 좌표 (intrinsic reaction coordinate) 경로로부터 규명하였다.

세 가지 방향족 아미노산의 구조 이성질체에 대해 M05-2X 는 그들의 실험적인 단열 이온화 에너지(adiabatic ionization energy)와 수직 이온화 에너지(vertical ionization energy)를 성공적으로 재현하는 반면, B3LYP 은 실험적으로 관찰된 값에 비해 0.2 ~ 0.3 eV 만큼 상당히 낮은 에너지를

산출하였다. 양이온 상태에서 그들의 구조 이성질체들에 대한 전하 분포를 자연적 결합 궤도 (natural bond orbital) 분석을 이용하여 조사하였다. 페닐알라닌의 구조 이성질체는 분자 내 수소결합의 형태에 따라 전하가 편재화 되거나 비편재화 되지만, 타이로신과 트립토판의 모든 구조 이성질체들은 그들이 이온화 되었을 때 분자 내 수소 결합의 형태와 무관하게 전하 편재화만 일어난다는 사실을 밝혀냈다.

세 가지 방향족 아미노산 가운데 오직 페닐알라닌에 대해서만 전하 비편재화가 일어나는 원리를 규명하기 위하여 간단한 이론 모델을 제시하였다. 그 모델의 기본 개념은 이들 방향족 아미노산이 내부적으로 두 개의 별도의 양이온 코어를 갖는 골격과 잔기로 구성된다는 가정에서 출발한다. 두 개의 양이온 코어 사이의 단열적 이온화 에너지의 차이가 전하 편재화의 정도를 결정한다는 사실을 확인하였다.

세 가지 방향족 아미노산은 생체 분자 내에서 발색단 역할을 하기 때문에 이들 분자의 전하 분포에 대한 분석 결과는 단백질이나 다른 아미노산에서의 전하 동력학을 해석하는데 있어서 중요한 단서를 제공할 것이다.

본 연구에서 다뤄진 세 가지 방향족 아미노산의 구조-의존적 특성과 이온화 에너지 사이에 밀접한 관련이 있음을 이론적으로 규명하였다.

주요어: M05-2X, 비공유결합 원거리 효과, 페닐알라닌, 타이로신과 트립토판, 전하 편재화, 이온화 에너지

학번: 2002-30939